

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/320407457>

Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis

Book · January 2007

CITATIONS

14

READS

6,527

5 authors, including:



Regina Kunz

University of Basel / University Hospital Basel

224 PUBLICATIONS 54,225 CITATIONS

[SEE PROFILE](#)



Günter Ollenschläger

University of Cologne

459 PUBLICATIONS 3,064 CITATIONS

[SEE PROFILE](#)



Heiner Raspe

Universitätsklinikum Münster

542 PUBLICATIONS 10,979 CITATIONS

[SEE PROFILE](#)



Günther Jonitz

Danube Private University

54 PUBLICATIONS 201 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Thoracic Pain in the Community (TOPIC) [View project](#)



Evidenzbasierte Medizin [View project](#)

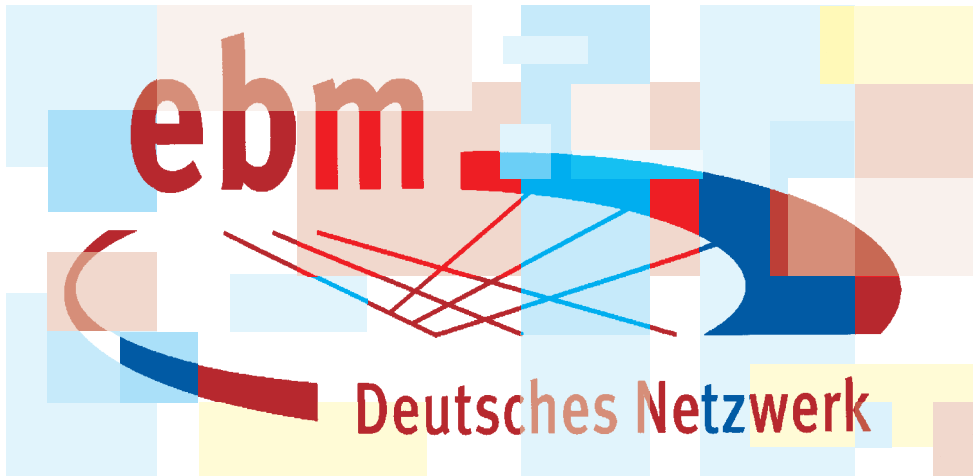
2. überarbeitete und erweiterte Auflage

R. Kunz / G. Ollenschläger / H. Raspe
G. Jonitz / N. Donner-Banzhoff (Hrsg.)

Lehrbuch Evidenz- basierte Medizin

in Klinik und Praxis

Koordination und Redaktion Marga Cox



Deutscher
Ärzte-Verlag

R. Kunz / G. Ollenschläger / H. Raspe / G. Jonitz /
N. Donner-Banzhoff (Hrsg.)
Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis

**R. Kunz / G. Ollenschläger / H. Raspe
G. Jonitz / N. Donner-Banzhoff (Hrsg.)**

Lehrbuch Evidenz- basierte Medizin

in Klinik und Praxis

Koordination und Redaktion Marga Cox

2. überarbeitete und erweiterte Auflage

Mit 60 Abbildungen und 85 Tabellen

Deutscher Ärzte-Verlag Köln

1. Auflage 2000

ISBN 978-3-7691-0538-4

aerzteverlag.de

Bibliografische Information Der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- oder Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Wichtiger Hinweis:

Die Medizin und das Gesundheitswesen unterliegen einem fortwährenden Entwicklungsprozess, sodass alle Angaben immer nur dem Wissensstand zum Zeitpunkt der Drucklegung entsprechen können.

Die angegebenen Empfehlungen wurden von Verfassern und Verlag mit größtmöglicher Sorgfalt erarbeitet und geprüft. Trotz sorgfältiger Manuskripterstellung und Korrektur des Satzes können Fehler nicht ausgeschlossen werden.

Der Benutzer ist aufgefordert, zur Auswahl sowie Dosierung von Medikamenten die Beipackzettel und Fachinformationen der Hersteller zur Kontrolle heranzuziehen und im Zweifelsfall einen Spezialisten zu konsultieren.

Der Benutzer selbst bleibt verantwortlich für jede diagnostische und therapeutische Applikation, Medikation und Dosierung.

Verfasser und Verlag übernehmen infolgedessen keine Verantwortung und keine daraus folgende oder sonstige Haftung für Schäden, die auf irgendeine Art aus der Benutzung der in dem Werk enthaltenen Informationen oder Teilen davon entstehen.

Das Werk ist urheberrechtlich geschützt. Jede Verwertung in anderen als den gesetzlich zugelassenen Fällen bedarf deshalb der vorherigen schriftlichen Genehmigung des Verlages.

Copyright © 2007 by
Deutscher Ärzte-Verlag GmbH
Dieselstraße 2, 50859 Köln

Umschlagkonzeption: Hans Peter Willberg und Ursula Steinhoff
Titelgrafik: Eva Kroll
Satz: Plaumann, 47807 Krefeld
Druck/Bindung: Zimmermann Druck, 58802 Balve

5 4 3 2 1 0 / 614

Vorwort

Gut sein wollen wir alle. Früher war das einfacher, zumindest, was den Faktor „Wissen“ angeht. Krankheiten waren noch Krankheiten und nicht „Zustände“ Therapien waren auf konkrete, vom Patienten erlebte Symptome und Diagnosen bezogen und nicht selten kausal und überschaubar. Diabetes mellitus zum Beispiel war durch Diät, dann durch Insulin, später durch die erste Generation oraler Antidiabetika behandelbar. Alles überschaubar, verständlich und praxisorientiert in der Anwendung.

Diese Zeiten sind vorbei. Die Medizin hat Fortschritte gemacht. Unsere Welt ist wesentlich komplexer geworden. Diese Fortschritte sind zum Teil enorm. Zum Beispiel in der Therapie von Malignomen oder in der Behandlung von Volkskrankheiten wie dem Herzinfarkt, wo neben der Akuttherapie auch die Bedeutung der konsequenten medikamentösen Nachbehandlung Standard geworden ist. Das Arsenal der Möglichkeiten, um Diabetiker „gesund“ zu halten, ist enorm.

Doch wie kommt das richtige Wissen in die Welt des praktizierenden Arztes? Neue Erkenntnisse zur richtigen Diagnostik und Therapie von Krankheiten werden fast nur noch durch groß angelegte Studien an großen Patientenkollektiven gewonnen, die jede für sich Möglichkeiten des Irrtums, der Verzerrung und der Manipulation aufweisen. Wert oder Unwert solcher Studien erschließen sich fast nur noch dem Experten. Unsere Wege, Wissen zu generieren und zu verbreiten, haben mit der aktuellen Medizin nicht Schritt gehalten. Studien weisen oft methodische Probleme auf; ihre Ergebnisse

werden nicht selten im Interesse der Pharmaindustrie mit großem Elan auf den Markt gebracht. Immer öfter stellt sich im Nachhinein heraus, dass der große, dem Arzt und den Patienten versprochene Fortschritt doch nur ein kleiner war und einer, der mit erheblichen Nebenwirkungen behaftet ist. Dies belegen die Erfahrungen mit COX2-Antagonisten oder die jüngere Diskussion über die Drug diluting stents bei der invasiven Kardiologie.

Auch das Vertrauen in Institutionen oder so genannte Meinungsbildner zur Übernahme von Wissen ist trügerisch. Selbst höchste angesehene wissenschaftliche Verlage knicken vor der Macht der Industrie ein und schreiten zur Zensur wohlbegründeter, kritischer Artikel.

Wer hilft? Was hilft? Antwort: ein neuer, kritischer und selbstkritischer Denkansatz mit Namen **Evidenzbasierte Medizin (EbM)**. In unübersichtlichen Zeiten frage man sich:

- ▲ Was braucht mein Patient – eigentlich?
- ▲ Was kann, was weiß ich – eigentlich? und
- ▲ Was sollte ich können und wissen – eigentlich?

Technik und Tugend der evidenzbasierten Medizin helfen dem Arzt, frühzeitig zu erkennen, was ihm zur richtigen Entscheidung und dem Patienten zum Wohle dient. Das vorliegende, umfassende Buch zur evidenzbasierten Medizin gibt die Möglichkeit, seine eigenen Fähigkeiten und Kenntnisse auf den neuesten, besten Stand zu bringen und besser zu differenzieren, ob ein neues, angepriesenes Verfahren Marketing oder tatsächlich Fortschritt bedeutet. Ebenfalls hilft es zu er-

kennen, was tatsächlich EbM bedeutet, nämlich keineswegs die Lehre von der doppelblind-randomisierten Studie, sondern die Suche nach der bestmöglichen Wahrheit, wie wir sie im Augenblick erkennen, auch bei miserabler Studienlage. Nachdem wir alle selbstverständlich neue Medikamente oder neue medizinische Techniken annehmen, haben wir hier die Chance, durch einen neuen Denk- und Handlungsansatz eine bessere Medizin zu betreiben.

Ein Lehrbuch, in dem ausgewiesene Kenner und Könner von Therapie und Praxis der Patientenversorgung ein so breites Spektrum wichtiger Bereiche der Medizin abdecken, ist selten. Hier ist es realisiert. Den Autoren herzlichen Dank, den Lesern und Nutznießern herzlichen Glückwunsch und herzlich willkommen in einer Medizin, die sich nicht einfach auf althergebrachte Methoden und Kenntnisse verlässt, sondern eigenständig in der Lage ist, den komplexen Dschungel des medizinischen Know-Hows zu durchforsten.

Selbst zu denken ist am Anfang anspruchsvoll. Gleichwohl ist es machbar (der Verfasser dieser Zeilen ist Chirurg und hat für sich die Grundlagen des Critical Appraisal begriffen und wieder viel Freude an seiner Arbeit gewonnen). Die selbstkritische Reflexion der Ergebnisse der eigenen Arbeit, wie

es die EbM-Philosophie vorgibt, führt zu mancherlei Aha-Erlebnissen und befruchtet den ärztlichen Alltag. Es führt in ein neues Zeitalter der medizinischen Aufklärung, das sich mit großer Expertise um Wahrheit, zumindest um mehr Wahrhaftigkeit und Transparenz in der Medizin bemüht. Es macht den Arzt wieder souverän und die Medizin glaubwürdiger und verlässlicher. So, wie es – eigentlich – sein sollte.

Dr. med. Günther Jonitz

Groß Glienicke bei Berlin, 15. Oktober 2006

Evidenzbasierte Medizin ist der Ausgang des **Arztes** aus seiner selbstverschuldeten Unmündigkeit.

Unmündigkeit ist das Unvermögen, sich seines Verstandes ohne die Hilfe eines anderen zu bedienen.

Selbstverschuldet ist diese Unmündigkeit, wenn die Ursache derselben nicht am Mangel des Verstandes, sondern der Entschließung und des Mutes liegt, sich seiner ohne die Leitung eines anderen zu bedienen.

Sapere aude! Habe den Mut, dich deines eigenen Verstandes zu bedienen!

Nach Immanuel Kant, Dezember 1784

Preface

No clinician would consider entering clinical practice without knowing the rudiments of history and physical examination. Nor would clinicians consider independent practice without a basic understanding of how the drugs they prescribe act on their patients. Yet, traditionally, clinicians have entered practice without an ability to understand evidence about how they should interpret what they find on history and physical examination, or the magnitude of the effects they might expect when they offer patients medication.

Evidence-based medicine provides a remedy for this problem. The movement to teach clinicians to become effective users of the medical literature began in the 1970's and evolved through the 1980s into a whole system for the delivery of clinical care. We needed a name for this new way of practice and the term „evidence-based medicine“, which first appeared in the medical literature in 1991 [1], proved extremely popular. Over the subsequent 15 years evidence-based medicine has evolved, and now represents not only an approach to effectively using the medical literature, but a principled guide for the process of clinical decision-making.

Members of the general public are surprised, and often appalled, when they learn that most physicians remain unable to critically read an original research article, or to fully understand the results reported there. For the physician, inability to critically appraise a research study and grasp all that is implied in its findings, limits their independence. The result is reliance on expert opinion, the practices of colleagues, and on

information from the pharmaceutical industry. But what is to do if experts and colleagues disagree, or if one is mistrustful of the enthusiastic advice from a pharmaceutical industry representative?

This book represents the key to a world that provides the answer to that question, a world that has traditionally been closed to most practicing physicians – the world of the original medical literature. Opening the door to this world is enormously empowering. No longer must one choose what to believe on the basis of which recommendation is backed up by the most authority, or speaks with the loudest voice. The ability to differentiate strong evidence from weak, and large treatment effects from small, allows clinicians to make independent judgements about what is best for their patients. It also allows them to explain the impact of alternatives to the patients themselves, and thus to ensure that choices are consistent with patients' underlying values and preferences.

Ten years ago, experts and the official voices of the organizations to which they belonged consistently recommended long-term hormone replacement therapy (HRT) for post-menopausal women. These recommendations were made largely on the basis of observational studies suggesting that women taking HRT could expect large reductions in their risk of major cardiovascular events. Proponents of evidence-based medicine raised concerns about the wisdom of strong advocacy of therapy for huge populations on the basis of the fundamentally weak methods of observational studies. Their voices were largely ignored, until randomized

trials demonstrated that the results of the observational studies were incorrect. If HRT has any impact on cardiovascular disease at all, it is to increase its frequency.

Many clinical communities now endorse widespread population screening to prevent the occurrence of cancer and cardiovascular disease. Breast cancer screening for women as young as 40 years, colon cancer screening for entire populations, and treatment to improve lipid profiles even in very low risk patients are widely advocated. Many clinicians are unaware that hundreds of individuals must be screened for breast or colon cancer, or treated with lipid-profile-modifying agents, for periods of up to a decade, to prolong a single life. The costs include anxiety as a result of the many false positive results, complications of invasive procedures such as lumpectomy or colonoscopy, side effects of treatment (including deaths as a result of a lipid-lowering agent now withdrawn from the market), and resource investment that, at least for some individuals, might be better allocated elsewhere.

The point is not that the experts were uniformly wrong in suggesting that women consider HRT, nor that screening or treatment of low-risk individuals to modify their cancer or coronary risk is wrong. Rather, clinicians should be aware there are important trade-offs in these decisions. If clinicians don't know the difference between an observational study and a randomized trial, or between a relative risk reduction and a risk difference, they are in no position to understand these trade-offs. If they are unable to understand the trade-offs, it is not possible for them to convey the possible benefits and risks to their patients, many of whom may, with a full understanding, decline screening or treatment.

This book provides the basic tools for the German speaking health care provider to evaluate the strength of original studies, to understand their results, and to apply those

results in day-to-day patient care. Increasingly, this does not mean a detailed reading of a large number of original studies. Rather, the caregiver can recognize valid pre-appraised resources, and differentiate them from poorly substantiated opinion. The book provides guides for assessing not only original studies of diagnosis and therapy, but systematic reviews that summarize a number of such original studies, and practice guidelines that summarize a whole body of evidence and take the leap from evidence to suggested action. Ability to differentiate strong from weak literature reviews, and to understand summaries of the magnitude of treatment effects, are crucial for efficient evidence-based practice. Ability to understand the strength of evidence behind a practice guideline, and the values and preferences implicit in recommendations, allows health care experts to efficiently incorporate recommendations into their daily practice.

When a new pivotal study comes along, evidence-based health care providers do not need to read it in detail to evaluate its significance, or to decide how to use its results. Imagine that I am telling you about a recently conducted study that reports an apparently important treatment effect. I tell you that the study was a randomized trial and that randomization was adequately concealed; that patients, caregivers, and those collecting and adjudicating outcome data were blind to whether patients received treatment or control interventions; that investigators successfully followed all patients who were randomized; and that, in the analysis, all patients were included in the groups in the groups to which they were randomized. Assuming that I am skilled in making these judgements, and am telling you the truth, you have all the information you need to judge the study's validity. If I then provide you with a few crucial details of who was enrolled, how the interventions were administered, and the magnitude and precision of

estimates of the intervention's impact on all patient-relevant outcomes, you have everything you need to apply the results in health care.

Synopses of individual studies that provide the crucial information needed to understand the appropriate strength of inference and apply the results are increasingly available. So are systematic reviews and, to a lesser extent, high quality evidence-based practice guidelines. Whole systems of knowledge based on evidence-based principles are beginning to arrive. When you, as a German health care provider, have read and digested the current text, you will have the tools to read and interpret synopses, systematic reviews, and evidence-based guidelines and textbooks, and be able to find such pearls in the rocky landscape of the current medical literature.

The authors and editors of this book are an accomplished group, deeply skilled in the principles and practice of evidence-based medicine. I am sorry to say I cannot vouch for the quality of the German text through my reading it in German. However, I can tell you that I have participated in workshop and seminars focused on EBM with a number of the authors, and have noted a profound understanding of evidence-based practice, and great skill in communicating its principles to audiences. They have provided a comprehensive and definitive text that will teach you to effectively search the medical literature and evaluate what you dis-

cover. In this text you will find case studies and examples directly relevant to your area of clinical practice. In addition, if you are an educator, you will find a whole section of the book devoted to issues relevant to this aspect of your vocation. If you are interested in the organization of health care, you will find an extensive discussion of how evidence bears on broader issues of health care delivery beyond the individual patient. Most important, you will find true-to-life examples of how to more effectively address the patient dilemmas you face in day-to-day care. In the end, you will find health care practice more satisfying and, most important, you will be more confident in providing your patients with optimal medical care. Finally, if you are interested in a deeper understanding of EBM, this book provides a stepping stone to a more comprehensive text that can provide knowledge and skills required for not only the practice, but also the teaching of EBM [2].

Prof. Gordon Guyatt MD, MSc
Hamilton, Kanada

Literatur

- [1] Guyatt GH, Evidence-based medicine. ACP J Club. (1991), 114, A-16
- [2] Guyatt G, Rennie D (eds.) (2002) Users' guides to the medical literature. A manual for evidence-based clinical practice. AMA Pr., Chicago

Inhaltsverzeichnis

I	Der Bedarf für EbM	1
1	Der Zweifel als Triebkraft des Erkenntnisgewinns in der Medizin	3
	<i>Johannes Köbberling</i>	
	1.1 Evidenz und Zweifel – ein Gegensatzpaar? – 3	
	1.2 Der Zweifel als grundlegende Methode der Wissenschaft – 4	
	1.3 Die Bedeutung des fehlenden oder unterdrückten Zweifels – 4	
	1.4 Gründe für fehlenden oder unterdrückten Zweifel bei der Gewinnung von wissenschaftlichen Erkenntnissen – 9	
	1.5 Mangelnder Zweifel beim Umgang mit wissenschaftlichen Erkenntnissen – 10	
	1.6 Des Kaisers neue Kleider – 12	
	1.7 Schlussfolgerung – 13	
2	Theorie, Geschichte und Ethik der Evidenzbasierten Medizin (EbM)	15
	<i>Heiner Raspe</i>	
	2.1 Zugänge zur EbM – 15	
	2.2 Evidenzbasierte Praxis der Medizin – 15	
	2.3 EbM und klinische Forschung – 16	
	2.4 EbM und die Theorie der Medizin – 18	
	2.5 EbM in medizinhistorischer Sicht – 20	
	2.6 Evidenzbasierte Medizin im Kontext aktueller Selbstverpflichtungen der ärztlichen Profession – 22	
	2.7 Ethik der Evidenzbasierten Medizin – 24	
	2.8 Schlussbemerkung – 28	
3	Evidenzbasierte Medizin macht Spaß oder: Das Richtige richtig machen	31
	<i>Wolfgang A. Blank und Günther Egidi</i>	
	3.1 Vorsicht: EbM kann süchtig machen! – 31	
	3.2 EbM in der Hausarzt-Praxis – Szenario – 31	
	3.3 Im Trott gefangen? – 31	
	3.4 Hilfsmittel EbM – 32	
	3.5 Viele Vorteile – 32	
	3.6 Das Werkzeug der EbM – 34	
	3.7 Zahlenspiele – 34	
	3.8 Wer profitiert? – 36	
	3.9 Sinnvolle Ergebnisse? – 37	
	3.10 Nachhaltige Zufriedenheit mit der eigenen Arbeit – kein Traum – 37	

4	Welche Evidenz braucht der Arzt?	39
	<i>Thomas Kühlein, Johannes Forster</i>	
4.1	Es gibt mehr als eine Evidenz – 39	
4.2	Validität externer Evidenz – 40	
4.3	Der Arzt braucht externe Evidenz, die eine konkrete Frage aus der Situation eines individuellen Patienten beantworten hilft – 41	
4.4	Der Arzt braucht schnell verfügbare Evidenz, der er vertrauen kann – 41	
4.5	Verschiedene Ärzte benötigen verschiedene Evidenz – 42	
4.6	Parameter der Relevanz – 45	
4.7	Die Indikation einer präventiven Therapie – 46	
4.8	Die Akzeptanz einer Prävention – 47	
4.9	Probleme im Alltag – 47	
4.10	Leitlinien im Alltag – 48	
4.11	Zusammenfassung – 48	
5	Welche Evidenz will der Patient?	51
	<i>Sylvia Sanger, Bernd Quadder, Frank Brunsmann</i>	
5.1	Einführung – 51	
5.2	Facetten der Evidenz in der Arzt-Patienten-Beziehung – 52	
5.3	Patientenerfahrungen – auch eine Evidenzquelle für die Medizin – 57	
5.4	Zusammenfassung – 57	
5.5	Werkzeugkasten für die Arzt-Patienten-Kommunikation – 58	
6	Welche Evidenz braucht das System?	61
	<i>Reinhard Busse, Bernhard Gibis</i>	
6.1	Methodische Überlegungen bezüglich evidenzbasierter Systementscheidungen – 61	
6.2	Bevölkerungs- bzw. systembezogene Effektivität: viele Faktoren und ein Modell – 64	
6.3	Ökonomische Daten für evidenzbasierte Systementscheidungen – 68	
6.4	Fallstudie Mammographie-Screening – 69	
7	Pluralismus in der Evaluation	75
	<i>Karl Wegscheider</i>	
7.1	Quellen der Evidenz – 75	
7.2	Die Methodik der Evidenzgenerierung – 77	
7.3	Therapieevaluation in einer pluralen Wissenschaftswelt – 84	
II	Die Technik und das Handwerkzeug	87
8	Vom Problem zur Frage	89
	<i>Regina Kunz</i>	
9	Kritische Bewertung von Studien zu Prävalenz und Symptomen	93
	<i>Andreas C. Sönnichsen</i>	
9.1	Klinisches Szenario – 93	
9.2	Die Evidenz suchen – 94	
9.3	Die Evidenz bewerten – kritische Evaluation einer Studie zur Prävalenz, Differenzialdiagnose und Prognose eines Symptoms – 95	
9.4	Auflösung des klinischen Szenarios – 99	

10a	Kritische Bewertung von Studien zur Ätiologie	101
	<i>Torsten Schäfer</i>	
10a.1	Klinisches Szenario – 101	
10a.2	Fall-Kontroll-Studie – 101	
10a.3	Kohortenstudien – 104	
10a.4	Die Evidenz – 104	
10a.5	Die Fall-Kontroll-Studie – 105	
10a.6	Die Kohortenstudie – 109	
10a.7	Fazit – 112	
10b	Genetische Studien	115
	<i>Inke R. König, Andreas Ziegler</i>	
10b.1	Die Evidenz bewerten – Kritische Bewertung einer genetisch-epidemiologischen Assoziationsstudie – 116	
11	Kritische Bewertung von Studien zu diagnostischen Tests	121
	<i>Heiner C. Bucher</i>	
11.1	Klinisches Szenario – 121	
11.2	Die Evidenz suchen – 122	
11.3	Die Evidenz bewerten – Kritische Bewertung einer Studie zur diagnostischen Testgenauigkeit – 122	
11.4	Auflösung des klinischen Szenarios – 130	
12	Kritische Bewertung von Studien zu therapeutischen Interventionen	133
	<i>Hilke Bertelsmann, Guido Lerzynski, Regina Kunz</i>	
12.1	Klinisches Szenario – 133	
12.2	Die Evidenz suchen – 133	
12.3	Die Evidenz bewerten – 134	
12.4	Was bedeuten die Ergebnisse der Studie? – 140	
12.5	Zusammenfassung der evidenzbasierten Studienbewertung – 145	
13	Kritische Bewertung von systematischen Reviews und Meta-Analysen	149
	<i>Heiner C. Bucher</i>	
13.1	Klinisches Szenario – 150	
13.2	Die Evidenz suchen – 150	
13.3	Die Evidenz bewerten – 151	
13.4	Auflösung des klinischen Szenarios – 157	
14	Kritische Bewertung von qualitativen Studien	159
	<i>Thorsten Meyer</i>	
14.1	Klinisches Szenario – 160	
14.2	Die Evidenz suchen – 160	
14.3	Die Evidenz bewerten – Kritische Bewertung einer qualitativen Studie – 162	
14.4	Sind die Ergebnisse der Studie glaubwürdig? – 167	
14.5	Wie lauten die Studienergebnisse? – 170	
14.6	Sind die Ergebnisse für die Behandlung meiner Patienten nützlich? – 172	
14.7	Auflösung des klinischen Szenarios – 174	
14.8	Ausblick – 175	

15	Kritische Bewertung von Leitlinien	177
	<i>Henning Thole, Frank Thalau, Günter Ollenschläger, Ina Kopp, Monika Lelgemann</i>	
15.1	Einleitung –	177
15.2	Das Szenario –	177
15.3	Qualitätskriterien von Leitlinien –	178
15.4	Bewertung formaler und methodischer Aspekte von Leitlinien –	178
15.5	Bewertung inhaltlicher Aspekte von Leitlinien –	178
15.6	Anwendung von DELBI –	181
15.7	Clearingverfahren –	182
15.8	Auflösung des Szenarios –	183
15.9	Zusammenfassung –	183
15.10	Ausblick: Nationale und internationale Aktivitäten zur Förderung der Leitlinienqualität –	184
16	Kritische Bewertung von Health Technology Assessment-Berichten	191
	<i>Matthias Perleth, Dagmar Lühmann</i>	
16.1	Einleitung –	191
16.2	Szenario –	193
16.3	HTA in Deutschland –	199
16.4	Internationale HTA-Aktivitäten –	199
17	Kritische Bewertung von gesundheitsökonomischen Studien	203
	<i>Reiner Leidl</i>	
17.1	Ein Szenario –	203
17.2	Warum die Wirtschaftlichkeit medizinischer Leistungen untersuchen? –	203
17.3	Wo setzt die ökonomische Evaluation an? –	204
17.4	Welche Standards gibt es für ökonomische Evaluationen? –	204
17.5	Welche Arten der ökonomischen Evaluation gibt es? –	205
17.6	Wie werden die Kosten gemessen? –	206
17.7	Wie wird das gesundheitliche Ergebnis gemessen? –	208
17.8	Was ist das Ergebnis der ökonomischen Evaluation? –	210
17.9	Kann Unsicherheit berücksichtigt werden? –	211
17.10	Was ist ökonomische Evidenz? –	211
17.11	Kann ökonomische Evidenz helfen, den Leistungskatalog zu gestalten? –	212
17.12	Welche anderen Entscheidungen kann ökonomische Evidenz unterstützen? –	213
17.13	Ausblick und Auflösung des Szenarios –	214
18	Kritische Bewertung von Gesundheitsinformationen für medizinische Laien	217
	<i>Sylvia Sängler, Marie-Luise Dierks</i>	
18.1	Hintergrund –	217
18.2	Szenarien –	218
18.3	Diskussion –	227
18.4	„Werkzeugkasten“ für Arzt und Patient –	229

19	Von der Evidenz zur Empfehlung	231
	<i>Regina Kunz, Monika Lelgemann, Gordon Guyatt, Gerd Antes, Yngve Falck-Ytter, Holger Schünemann</i>	
19.1	Die GRADE-Methodik: Mehr Transparenz bei der Erstellung von abgestuften Handlungsempfehlungen –	232
19.2	Ausgewählte Aspekte aus der Leitlinienarbeit in Deutschland –	241
20	Wo erhalte ich Antwort auf meine Fragen?	249
	<i>Monika Lelgemann, Norbert Donner-Banzhoff</i>	
20.1	Einleitung –	249
20.2	Eine schwierige Abwägung –	249
20.3	Systematische Aufbereitung: die Evidenzpyramide –	251
20.4	Der „Weg des Wissens“ –	252
20.5	Bias überall –	254
20.6	Zusammenfassung –	255
20.7	Ausblick –	255
III	EbM im ärztlichen Alltag	259
21	Ärztliches Denken und Entscheiden	261
	<i>Norbert Donner-Banzhoff</i>	
21.1	Szenario –	261
21.2	Das kognitive Kontinuum –	261
21.3	Der Sinn der Routine –	262
21.4	Diagnosen –	262
21.5	Faustregeln und Heuristiken –	262
21.6	EbM: Instrument der kritischen Reflexion –	263
21.7	EbM-Rezeption: anders als erwartet –	264
21.8	EbM – nicht jedem willkommen –	264
22	Evidenzbasierte Medizin im klinischen Alltag – Chirurgie	267
	<i>Stefan Sauerland, Karl-Heinz Moser, Edmund A. M. Neugebauer</i>	
22.1	Klinisches Szenario –	267
22.2	Medizinischer Hintergrund –	267
22.3	Die Evidenz –	267
22.4	Auflösung des Szenarios –	269
22.5	Kommentierung und Ausblick –	269
23	Medikamentöse Therapie des Morbus Alzheimer: Ein Disput um die Evidenz zwischen Studienkritik und Klinik	271
	<i>Hanna Kaduszkiewicz, Claus Wächtler</i>	
23.1	Hintergrund –	271
23.2	Zwei Patienten mit Alzheimer-Demenz –	271
23.3	Der Abwägungsprozess –	272
23.4	Kommentar –	275
23.5	Zusammenfassung –	275

24	Mammographie-Screening/Prävention – Kritische Bewertung des Screening/Patienteninformation	277
	<i>Ingrid Mühlhauser</i>	
24.1	Einleitung/Hintergrund –	277
24.2	Szenarien –	277
24.3	Kampagnen führen zu Trugschlüssen und Irrtümern –	277
24.4	Evidenzbasierte Informationen für eine informierte Entscheidung –	278
24.5	Einschätzung des individuellen Krebsrisikos –	278
24.6	Instrumente zur Vorhersage des individuellen Brustkrebsrisikos –	278
24.7	Einschätzung des möglichen Nutzens, Schadens bzw. Fehlens von Nutzen –	279
24.8	Auflösung der Szenarien –	279
24.9	Schlussfolgerungen –	280
25	EbM in der Rheumatologie – die interdisziplinäre Leitlinie: Management der frühen Rheumatoiden Arthritis	283
	<i>Monika Lelgemann, Matthias Schneider</i>	
25.1	Hintergrund –	283
25.2	Versorgungsproblem –	283
25.3	Klinisches Szenario – Teil A –	284
25.4	Leitlinien als Quelle aufbereiteter Evidenz –	284
25.5	Klinisches Szenario – Teil B –	286
25.6	Barrieren der Leitlinien-Anwendung –	286
25.7	Diskussion –	286
25.8	Auflösung des klinischen Szenarios –	287
26	Evidenzbasierte Medizin in der Allgemeinarztpraxis	289
	<i>Annette Becker, Norbert Donner-Banzhoff</i>	
26.1	Klinisches Szenario –	289
26.2	Die Evidenz suchen –	289
26.3	Die Evidenz bewerten –	290
26.4	Klinisches Szenario –	291
26.5	Diskussion –	292
27	Individualisierte Risikoprognose für die Herz-Kreislauf-Prävention: ein Werkstattbericht	295
	<i>Norbert Donner-Banzhoff, Heidemarie Keller, Eva-Maria Sadowski, Tanja Krones, Erika Baum, Andreas C. Sönnichsen, Uwe Popert</i>	
27.1	Die Beratungsstrategie „ARRIBA-Herz“ –	295
27.2	Entscheidungen –	296
27.3	Fazit –	302
28a	Revaskularisation bei koronarer Herzerkrankung	305
	<i>Martin Ruß, Karl Werdan</i>	
28a.1	Klinisches Szenario –	305
28a.2	Revaskularisation oder medikamentöse Therapie? –	305
28a.3	Klinisches Szenario –	307
28a.4	Revaskularisation: operativ oder interventionell? –	307
28a.5	Klinisches Szenario –	309
28a.6	Fazit –	309

28b	Chirurgische Revaskularisation bei koronarer Herzkrankheit	313
	<i>Robert von Wattenwyl, Hans-Reinhard Zerkowski</i>	
28b.1	Klinisches Szenario – Chirurgisches Konsilium –	313
28b.2	Bypassoperation oder medikamentöse Therapie? –	313
28b.3	Klinisches Szenario – Nachfrage des Patienten –	314
28b.4	Revaskularisation: Operation oder Katheterintervention? Hat die Operation Einfluss auf die Prognose? –	314
28b.5	Revaskularisation: Operation oder Katheterintervention? Ist bei immer besserem Material die wiederholte Katheterbehandlung nicht auch eine Option anstelle Operation? –	315
28b.6	Klinisches Szenario – Alter und Risiko –	316
28b.7	Alter und Risikoabschätzung bei chirurgischer Koronarrevaskularisation –	316
28b.8	Fazit des Herzchirurgen –	317
29	Evidenzbasierte Medizin am Beispiel des lumbalen Bandscheibenvorfalles	319
	<i>Josef Ramsbacher</i>	
29.1	Einleitende Übersicht –	319
29.2	Klinisches Szenario –	320
29.3	Kritische Analyse der vorhandenen Evidenz –	320
29.4	Diskussion –	321
29.5	Auflösung des klinischen Szenarios –	322
29.6	Schlussfolgerung –	322
30	Schlaganfall und Rehabilitation	325
	<i>Günter Seidel</i>	
30.1	Einleitung –	325
30.2	Klinisches Szenario –	325
30.3	Evidenz physiotherapeutischer rehabilitativer Therapieprinzipien –	325
30.4	Klinisches Szenario –	327
30.5	Evidenz pharmakologischer rehabilitativer Therapieprinzipien –	327
30.6	Klinisches Szenario –	329
31	Evidenzbasierte Medizin in der Pädiatrie – Beispiel Pseudokrapp	331
	<i>Johannes Forster, Dirk Bassler</i>	
31.1	Das Versorgungsproblem –	331
31.2	Klinisches Problem –	332
31.3	Festlegung der Fragestellung –	332
31.4	Die Evidenz suchen und bewerten –	332
31.5	Schlussfolgerung für die Praxis –	334
32	Evidenzbasierte Medizin bei psychischen Erkrankungen	335
	<i>Michael M. Berner, Mathias Berger, Martin Härter</i>	
32.1	EbM in der Psychiatrie und Psychotherapie –	335
32.2	Anwendungsbeispiel –	338
32.3	Schlussfolgerungen –	341

33	Evidenzbasierte Medizin in der Sozialmedizin: Beispiel eines Grundsatzgutachtens zur Vertebroplastie und Kyphoplastie	345
	<i>Johannes Giehl, Christoph Kreck</i>	
33.1	Einführung – 345	
33.2	Sozialmedizinischer Hintergrund und Versorgungsproblem – 345	
33.3	Auftrag und Arbeitsgruppe zur Gutachtenerstellung – 346	
33.4	Arbeitsplanung und Literaturrecherche – 347	
33.5	Darstellung der Evidenz – 347	
33.6	Synthese des Wissens – 349	
33.7	Sozialmedizinische Empfehlung – 349	
34	Evidenzbasierte Medizin in der Zahnmedizin – Beispiel Myoarthropathien des Kausystems	351
	<i>Jens C. Türp</i>	
34.1	Das Versorgungsproblem – 351	
34.2	Klinisches Szenario – 351	
34.3	Die Evidenz bewerten – 352	
34.4	Diskussion – 354	
34.5	Auflösung des klinischen Szenarios – 355	
34.6	Schlussfolgerungen – 356	
IV	EbM in den Strukturen der Gesundheitsversorgung	359
35	Evidenzbasierte Medizin und Leitlinien	361
	<i>Ina Kopp, Monika Lelgemann, Günter Ollenschläger</i>	
35.1	Was sind Leitlinien und wozu dienen sie? – 361	
35.2	Wie werden evidenz- und konsensbasierte Leitlinien entwickelt? – 363	
35.3	Redaktion und Verbreitung – 368	
35.4	Implementierung – 369	
35.5	Evaluierung und Fortschreibung – 370	
35.6	Schlussfolgerung – EbM und Leitlinien – 371	
36	Systemsteuerung im Rahmen des SGB V: der Gemeinsame Bundesausschuss	375
	<i>Bernhard Gibis, Regina Klakow-Franck, Nicole Schlottmann, Johannes Bruns</i>	
36.1	Entscheidungsfindung im Gemeinsamen Bundesausschuss – 376	
36.2	EBHC im G-BA – 379	
36.3	Erfahrungen mit der Anwendung der Prinzipien der EBHC im G-BA – 383	
36.4	Ausblick – 385	
37	Evidenzbasierte Medizin und ethische Aspekte	387
	<i>Nikola Biller-Andorno, Christian Lenk</i>	
38.1	EbM in der Forschung – 387	
38.2	EbM bei der Ressourcenallokation – 389	
38.3	EbM und Versorgungsqualität – 389	
38.4	Schlussfolgerungen – 390	

38a	Evidenzbasierte Medizin und Recht: Haftungsrecht	393
	<i>Dieter Hart</i>	
	38a.1 Einleitung – 393	
	38a.2 Haftungsrecht als Verhaltensnormenrecht – 393	
	38a.3 Schnittstellen von EbM und Haftungsrecht – 394	
	38a.4 Standard, Leitlinien und EbM – 396	
	38a.5 Folgerungen für gerichtliche Behandlungsfehlerverfahren – 399	
38b	Evidenzbasierte Medizin und Recht: Sozialrecht	401
	<i>Felix Welti</i>	
	38b.1 Einführung – 401	
	38b.2 Gesetzliche Grundlagen, insbesondere in der gesetzlichen Krankenversicherung (SGB V) – 401	
	38b.3 Verfahrensordnung des G-BA – 405	
	38b.4 Rechtsprechung der Sozialgerichtsbarkeit – 407	
	38b.5 Die Entscheidung des Bundesverfassungsgerichts vom 6.12.2005 – 407	
	38b.6 Schluss – 409	
V	Aus-, Weiter- und Fortbildung	411
39	Ausbildung	413
	<i>Martin Bergold, Tobias Weberschock</i>	
	39.1 Sie können unterscheiden zwischen evidenzbasierter Lehre und Lehre der EbM – 413	
	39.2 Sie können 2 verschiedene Szenarien zur schrittweisen Ausbildung beschreiben – 414	
	39.3 Sie können die Besonderheiten der schrittweisen Ausbildung in EbM erklären – 414	
	39.4 Sie können 12 Tipps für einen gelungeneren Unterricht aufzählen – 417	
	39.5 Sie können die Eigenschaften von formativen und summativen Evaluationen definieren – 418	
	39.6 Stand der Ausbildung und Ausblick – 419	
40	Weiter- und Fortbildung	421
	<i>Norbert Donner-Banzhoff, Olaf Weingart, Gerd Burmester, Susanne Weinbrenner, Hans Wille, Regina Kunz</i>	
	40.1 EbM: Von der Etablierung zur Routine – 421	
	40.2 Didaktische Formen – 422	
	40.3 Das Curriculum Evidenzbasierte Medizin der Bundesärztekammer und des DNEbM e. V. – 424	
	40.4 EbM-International: Leonardo da Vinci – 425	
	40.5 Die EbM-Zukunft – 426	

41	Schulung von medizinischen Laien	427
	<i>Bettina Berger, Nicole Skoetz, Olaf Weingart</i>	
41.1	Patientenschulung in EbM – ein reales Szenario –	427
41.2	EbM-Training für Laien –	427
41.3	Warum EbM-Training für Laien? –	427
41.4	Projekte zur EbM-Laienschulung in Deutschland – mehr als ein theoretisches Modell –	429
41.5	Ergebnisse –	432
41.6	Ziele in der Zukunft –	434
VI	Deutsche EbM-Organisationen	437
42	Das Deutsche Netzwerk Evidenzbasierte Medizin (DNEbM)	439
	<i>Jürgen Windeler</i>	
42.1	Zielsetzungen –	439
42.2	Struktur –	440
42.3	Leistungen –	440
43	Cochrane Collaboration: Deutsches Cochrane Zentrum/ Deutsche Cochrane-Review-Gruppen	443
	<i>Antje Timmer, Gerd Antes</i>	
43.1	Die internationale Cochrane Collaboration –	443
43.2	Die Cochrane Library –	447
43.3	Das Deutsche Cochrane Zentrum –	448
43.4	Die deutschen Cochrane-Review-Gruppen –	452
44	Das ÄZQ – Kompetenzzentrum der ärztlichen Selbstverwaltung für Qualität und Evidenz in der Medizin	457
	<i>Günter Ollenschläger, Christian Thomeczek, Sylvia Sänger, Monika Lelgemann</i>	
44.1	Einleitung –	457
44.2	Rahmenbedingungen –	457
44.3	Organisationsform –	458
45	Das Netzwerk der Koordinierungszentren für Klinische Studien	467
	<i>Christoph M. Seiler, Insa Bruns, Stephanie Wolff, Walter Lehmacher, Markus Löffler, Herbert Maier-Lenz</i>	
45.1	Hintergrund –	467
45.2	Aufgaben der Koordinierungszentren für klinische Studien –	469
45.3	Aus- und Weiterbildung im KKS-Netzwerk –	472
45.4	Zusammenfassung und Ausblick sowie künftige Herausforderungen –	474
46	DIMDI – Deutsches Institut für Medizinische Dokumentation und Information ...	475
	<i>Alric Rüther, Britta Göhlen</i>	
46.1	Medizininformation umfassend und unabhängig –	475
46.2	Informationssystem Gesundheitswesen –	475
46.3	Informationssystem Health Technology Assessment: Wissenschaft für Politik –	477
46.4	Informationssystem Arzneimittel –	479
46.5	Informationssystem Medizinprodukte –	479
46.6	Medizinische Klassifikationen und Begriffssysteme –	480

VII Ressourcen	481
47 Bias, Confounding, Chance	483
<i>Jürgen Windeler</i>	
47.1 Interne und externe Validität –	483
47.2 Bias –	484
47.3 Prinzip statistischer Aussagen –	485
47.4 Den Zufall im Griff –	486
48 Checklisten zur Bewertung von klinischen Studien und Leitlinien	491
<i>Guido Lerzynski, Hilke Bertelsmann, Regina Kunz</i>	
48.1 Bewertungsinstrumente, die primär für Autoren zur Verbesserung der Berichtsqualität von Studien entwickelt wurden –	492
48.2 Bewertungsinstrumente, die primär für die Nutzer von evidenzbasierten Studienergebnissen entwickelt wurden –	493
49 Glossar zur Evidenzbasierten Medizin	497
<i>Regina Kunz, Dagmar Lühmann, Jürgen Windeler, Monika Lelgemann, Norbert Donner-Banzhoff</i>	
VIII Anhang	511
Für Ärzte	513
Datenbankrecherche: Exemplarische Suche in Medline über PubMed – 513	
✂ Toolbox Evidenzquellen für den Arzt – 541	
✂ Toolbox Materialien für Patienten zur Dokumentation des Erkrankungsverlaufes – 542	
Für Patienten	000
✂ Toolbox Hilfestellung bei der Einschätzung der Qualität von Informationen im Internet – 543	
✂ Toolbox Verlässliche Informationsangebote im Internet – 544	
✂ Toolbox Unterstützung des Arzt-Patienten-Gesprächs – 545	
✂ Toolbox Erfahrungsaustausch mit ebenfalls Betroffenen – 546	
Herausgeber- und Autorenverzeichnis	547
Stichwortverzeichnis	555

I Der Bedarf für EbM

1	Der Zweifel als Triebkraft des Erkenntnisgewinns in der Medizin	3
	<i>Johannes Köbberling</i>	
2	Theorie, Geschichte und Ethik der Evidenzbasierten Medizin (EbM)	15
	<i>Heiner Raspe</i>	
3	Evidenzbasierte Medizin macht Spaß oder: Das Richtige richtig machen	31
	<i>Wolfgang A. Blank und Günther Egidi</i>	
4	Welche Evidenz braucht der Arzt?	39
	<i>Thomas Kühlein, Johannes Forster</i>	
5	Welche Evidenz will der Patient?	51
	<i>Sylvia Sänger, Bernd Quadder, Frank Brunsmann</i>	
6	Welche Evidenz braucht das System?	61
	<i>Reinhard Busse, Bernhard Gibis</i>	
7	Pluralismus in der Evaluation	75
	<i>Karl Wegscheider</i>	

1 Der Zweifel als Triebkraft des Erkenntnisgewinns in der Medizin

Johannes Köbberling

„Dinge zu bezweifeln, die ganz ohne weitere Untersuchung jetzt geglaubt werden, das ist die Hauptsache überall“

Georg Christoph Lichtenberg (1742–1799) (Dieses Zitat und alle weiteren den Kapiteln vorangestellten Aphorismen stammen von Georg Christoph Lichtenberg, dem Göttinger Physiker und geistreichen Aufklärer.)

1.1 Evidenz und Zweifel – ein Gegensatzpaar?

Es bedarf zumindest einer Begründung, wenn ein Lehrbuch der Evidenzbasierten Medizin mit einer Abhandlung über die große Bedeutung des Zweifels beginnt. Sollte eine nach gründlicher Suche endlich erzielte Evidenz nicht den Zweifel hinter sich lassen? Diese Frage muss verneint werden! Dabei ist allerdings vorauszusetzen, dass das Wort Evidenz anders aufgefasst wird, als es nach dem allgemeinen deutschen Sprachgebrauch üblich ist.

Schon oft wurde betont, dass das in der deutschen Sprache übliche Verständnis des Begriffes Evidenz für die Evidenzbasierte Medizin schlecht übertragbar ist. Aber auch das angelsächsische Verständnis von „evidence“ als bestmöglicher Beleg für die Richtigkeit einer Annahme muss ja nicht notwendigerweise den Zweifel einschließen. Die Denkweise, die zur Evidenzbasierten Medizin führt, hat aber den ständigen Zweifel als wesentliche Grundlage des Erkenntnisgewinns in der Medizin einbezogen. Ihr Leitbild ist das des kritikfähigen und jederzeit zum

Zweifel bereiten Arztes. So verwendet Sackett, der wohl bekannteste Protagonist von Evidence based medicine, das Wort „evidence“ nach seiner eigenen Aussage immer im Sinne Poppers, nämlich als eine Hypothese, die zur Falsifizierung offen ist. In diesem Zusammenhang sei auch ein Satz des Gießener Pharmakologen Habermann zitiert: „Evidence based medicine bringt ins Bewusstsein, wofür die Evidenz fehlt.“ Der Zweifel am Vorhandensein von Evidenzen kommt demnach aus derselben Grundhaltung wie die Suche nach vorhandenen Evidenzen. In diesem Sinne sind Zweifel und Evidenz also nicht Gegensätze, sondern eher die beiden Seiten der gleichen Medaille. Evidenzbasierte Medizin kann sich nur in einem Umfeld entwickeln, in dem der Zweifel jederzeit gepflegt und gefördert wird.

Nach der inzwischen weitgehend etablierten Methode besteht Evidence based medicine aus den folgenden Komponenten:

- ▲ Formulierung beantwortbarer klinischer Fragen,
- ▲ Suche nach der besten externen Evidenz,
- ▲ Kritische Bewertung dieser Evidenz bezüglich Validität und klinischer Relevanz,
- ▲ Umsetzung dieser Erkenntnisse in die klinische Arbeit,
- ▲ Bewertung der eigenen Leistung.

In diese häufig zitierte Liste sollte ein weiterer Punkt aufgenommen werden:

- ▲ Zweifel am bisherigen Vorgehen und an der Gültigkeit von Lehrmeinungen in der Medizin.

Mit den folgenden Ausführungen soll verdeutlicht werden, warum dem Zweifel eine so zentrale Rolle in der Denk- und Handlungsweise der Evidenzbasierten Medizin zukommt.

*„Wahrhaftes unaffektiertes Mißtrauen
in allen Stücken ist
das sicherste Zeichen von
Geistesstärke“*

1.2 Der Zweifel als grundlegende Methode der Wissenschaft

Die grundlegenden philosophischen Auseinandersetzungen mit dem Wissenschaftsbegriff in diesem Jahrhundert gehen auf Karl Popper zurück [1]. Nach ihm ist Wissenschaft nicht Gewissheit, auch nicht Suche nach Gewissheit. Die wissenschaftliche Erkenntnis besteht vielmehr in der permanenten Suche nach objektiv wahren, erklärenden Theorien. Diese Suche besteht darin, den Fehler, den Irrtum zu bekämpfen und alles zu tun, um Unwahrheiten zu entdecken und auszuschließen. Ausgehend von der sokratischen Einsicht in unser Nichtwissen hat er seine Fehlbarkeitslehre begründet. Statt von Wissen im Sinne von Gewissheit redet er von Vermutungswissen oder Theorien. Manche Theorien können wahr sein, aber auch wenn sie wahr sind, so können wir das niemals sicher wissen, weil es kein objektives Kriterium der Wahrheit gibt.

Es gibt nach Popper aber ein Kriterium der Wissenschaftlichkeit, nämlich die Bereitschaft zur ständigen kritischen Überprüfung und gegebenenfalls Verwerfung der Hypothesen. Der ständige Zweifel, der zu immer neuen Versuchen der Falsifikation führt, ist somit einer der wesentlichen Motoren für den wissenschaftlichen Erkenntnisgewinn. Wissenschaftlicher Fortschritt entsteht durch die Bemühung, immer feinere Siebe der Falsifikation zu konstruieren und dadurch zu

immer richtigeren Aussagen über unsere Welt zu gelangen. Vereinfacht ausgedrückt, liegt dem Erkenntnisgewinn die Schleife „Hypothese – Zweifel – Falsifikation – neue Hypothese“ zugrunde. Wenn in diesem fortlaufenden Prozess der Zweifel fehlt oder unterentwickelt ist, ist der Fortschritt behindert.

*„Lass Dich nicht anstecken, gib keines
anderen Meinung,
ehe Du sie Dir anpassend gefunden,
für Deine aus,
meine lieber selbst“*

1.3 Die Bedeutung des fehlenden oder unterdrückten Zweifels

Falsifikation und Formulierung einer neuen Hypothese sind manchmal sehr un bequem, schwierig oder aufwendig. Es gibt aber auch Beispiele, bei denen Falsifikation und Erstellung einer neuen Hypothese sehr leicht gewesen wären, bei denen der wissenschaftliche Stillstand ausschließlich durch den fehlenden Zweifel entstanden ist.

Sehr eindrucksvoll ist in diesem Zusammenhang folgende Geschichte: Fast zwei Jahrtausende lang wurde angenommen, dass die Frau weniger Zähne habe als der Mann, weil der berühmte Aristoteles dies so gelehrt hatte. Die entsprechende Behauptung von Aristoteles wäre extrem leicht zu widerlegen gewesen. Niemand hat aber die notwendige Falsifikation, die in einem einfachen Nachzählen der Zähne bestanden hätte, vorgenommen, weil man der berühmten Autorität glaubte und deshalb gar nicht erst zweifelte.

Die entscheidende Bedeutung des Zweifels im Prozess des Erkenntnisgewinns wird immer dann leicht erkennbar, wenn ein angemessener Zweifel, aus welchen Gründen auch immer, ausbleibt. Im Folgenden soll mit drei weiteren Beispielen das Problem illustriert und vertieft werden. Abschließend

soll dann eine Analyse erfolgen, wo die Gründe für fehlenden oder unterdrückten Zweifel bei der Aufstellung von Hypothesen durch den Wissenschaftler zu suchen sind, warum auch im Umgang mit wissenschaftlichen Erkenntnissen ein derart eklatanter Mangel an Zweifel besteht, und schließlich, was zur besseren Entwicklung des Zweifels in der Medizin getan werden kann.

*„Der Mensch ist ein solches
Wunder von Seltsamkeit,
dass ich überzeugt bin,
es gibt Leute, die meinen,
sie glaubten etwas und glaubens doch nicht,
die sich selbst belügen, ohne es zu wissen“*

1.3.1 Beispiel 1: Abderhalden und seine Abwehrfermente

Besonders exemplarisch und darüber hinaus zeitgeschichtlich sehr interessant ist in diesem Zusammenhang die Geschichte von Emil Abderhalden und seinen „Abwehrfermenten“, über die in der Zeitschrift *Nature* berichtet wurde [2]. Abderhalden, Jahrgang 1877, hatte sich schon sehr früh mit der Synthese von Eiweißstoffen und der Wirkung von Proteasen beschäftigt. Bereits 1908, also mit 31 Jahren, wurde er Professor für Physiologie an der Tierärztlichen Hochschule in Berlin und drei Jahre später Professor für Physiologie und physiologische Chemie an der Universität Halle, wo er von 1911 bis zu seinem Tode im Jahre 1950 wirkte. Im Jahre 1909 publizierte er zum ersten Mal über so genannte Schutzfermente, die er später Abwehrfermente nannte. Schon 1912 hat er hierüber ein ganzes Buch herausgegeben, das bereits 1914 in einer zweiten Auflage erschien.

Nach Abderhaldens Theorie produzieren Tiere und Menschen spezifische Proteasen, wenn sie mit Fremdeiweiß in Berührung kommen. Es sollte sich hierbei also um einen

ähnlichen Vorgang handeln wie die Bildung von Antikörpern, und man könnte die Abwehrfermente auch als eiweißabbauende spezifische Antikörper auffassen. Es handelt sich um eine bestechende Hypothese – nur ist sie leider falsch. Aus ihr entstand aber ein gigantisches Konstrukt aus Selbsttäuschung, Lüge und Betrug.

Die größte praktische Anwendung der vermeintlichen Abwehrfermente erfuhr die folgende „Entdeckung“. Unter der Vorstellung, dass die Plazenta, der Mutterkuchen, für die schwangere Frau ein Fremdeiweiß darstellt, sollten bei schwangeren Frauen Abwehrfermente gegen plazentare Proteine entstehen. Zum Zweck eines Tests wurden deshalb aus Plazentagewebe denaturierte unlösliche Eiweiße gewonnen. Wenn diese mit Serum von schwangeren Frauen in Berührung gebracht wurden, sollten durch die Einwirkung der Abwehrfermente Peptide im Serum entstehen, die sich herausdialysieren und anschließend durch eine einfache Eiweißreaktion nachweisen lassen. Blutserum von nicht schwangeren Frauen oder von Männern sollten eine solche Reaktion nicht zeigen.

Dieser Test hat verständlicherweise die Gynäkologen fasziniert, gab es bis dahin doch keinen zuverlässigen serologischen Schwangerschaftstest, und der Ultraschall, mit dem heute jede Schwangerschaft leicht erkannt werden kann, war noch lange nicht erfunden. In den Jahren 1912 und 1913 sind 25 wissenschaftliche Arbeiten aus verschiedenen gynäkologischen Laboratorien erschienen, die sich mit Abderhaldens Schwangerschaftstest beschäftigten, fast ausschließlich mit positivem Ergebnis. 1914 wurde eine Umfrage unter den Direktoren der deutschen Universitäts-Frauenkliniken bezüglich der persönlichen Erfahrungen mit diesen Tests durchgeführt. Unter den 15 Professoren, die antworteten, berichteten alle über mehr oder weniger positive Ergebnisse, niemand über negative Ergebnisse. Die Begeisterung stieg

daraufhin weiter. In der vierten Auflage des Buches über die Abwehrfermente aus dem Jahre 1914 hat Abderhalden schon 451 Arbeiten zitiert, die den Gebrauch seines Testes beschreiben, viele auch aus ausländischen Zeitschriften.

Ähnlich wie der Schwangerschaftstest sollte auch ein Test für den Nachweis von Sarkomen und anderen Karzinomen funktionieren. Auch Infektionskrankheiten wie die Syphilis sollten sich mit einem analogen Test nachweisen lassen. Ganz besondere Blüten trieb die Abderhalden'sche Reaktion in der Psychiatrie. In diesen Jahren kurz vor und kurz nach dem ersten Weltkrieg war man fieberhaft auf der Suche nach organischen Markern für psychische Erkrankungen. Während Abderhalden selbst seine Reaktion nur bei der Paralyse positiv fand, berichteten andere über Abwehrfermente im Blut von Schizophrenen oder von Dementia präcox-Kranken. Nachdem einige hundert Arbeiten zu diesen Themen erschienen waren, wurden große Feldforschungen durchgeführt, ohne dass je kritisch geprüft wurde, ob die Methode überhaupt das misst, was zu messen sie vorgab. Man war von dem Konzept derart überzeugt, dass ein Forscher schon im Jahre 1913 einmal äußerte, „paradoxe Reaktionen sind Ausfluss von Versuchsfehlern“.

Wenn aber doch Zweifler auftauchten, gab es ein ganzes Arsenal von wirksamen Abwehrmechanismen. Der später sehr bekannte Biochemiker Michaelis wurde vom Chef des Krankenhauses, in dessen Labor er in seinen jungen Jahren arbeitete, gebeten, die Aussagekraft des Abderhalden'schen Schwangerschaftstest zu überprüfen. Er stellte fest, dass die Experimente sich nicht reproduzieren ließen, obgleich er eine ganze Woche in Abderhaldens Labor zugebracht hatte, um die Methode genau zu erlernen. 1914 hat er diese negativen Ergebnisse publiziert, und dies führte zu dem Ende seiner wissenschaftlichen Karriere in Deutschland. Er hat daraufhin Deutschland verlassen und bekannt-

lich eine bedeutende wissenschaftliche Karriere in Amerika gemacht.

Abderhalden hat aber bis zu seinem Tode 1950 weiter über die Abwehrfermente publiziert. Seine Strategie war sehr einfach. Er brauchte Mitarbeiter, die das fanden, was sie für ihn finden sollten. Er hat immer wieder argumentiert, dass der Schwangerschaftstest und viele seiner anderen Tests bei einer großen Zahl von Laboratorien einwandfrei gewirkt haben, und dass sich eine derart große Zahl von Wissenschaftlern einfach nicht täuschen kann.

Noch in den 30er- und 40er-Jahren wurden in vielen deutschen Instituten die Abwehrfermente im Zusammenhang mit sehr unterschiedlichen Fragestellungen bearbeitet. So forschte man nicht nur weiter über Krebstests, sondern auch über mögliche Krebstherapien. Auch bei den Infektionskrankheiten erweiterte man die Studien zur Diagnostik auf solche zur Therapie. Sogar die in dieser Zeit modernen verschiedenen psychologischen Typen nach Kretschmar sollten über Abwehrfermente unterschieden werden können. Nach dem Tode Abderhaldens hat sein Sohn die Forschungen fortgesetzt und behauptet, dass über die Abwehrfermente eine perfekte diagnostische Möglichkeit bestände, die optimalen Zelltypen bei der obskuren Frischzellentherapie nach Niehans festzulegen. Spätestens hier wird erkennbar, dass die entsprechende „Forschung“ den Boden der Wissenschaft endgültig verlassen hat.

Erwähnt werden muss noch eine schlimme Verbindung dieser Forschung zur kriminellen Naziforschung. Im Jahre 1942 hat der Humangenetiker Ottmar von Verschür, der übrigens auch nach dem Krieg noch in Münster als Ordinarius tätig war, mit seinem früheren Assistenten Josef Mengele zusammen einen Antrag formuliert, um in Auschwitz über die Abwehrfermente zu forschen, die bei Menschen unterschiedlicher Rassen nach einer absichtlichen Infektion mit ver-

schiedenen Krankheiten entstehen. Mengele hat über 200 Blutproben von künstlich infizierten Juden und Sinti und Roma, darunter einigen Zwillingspaaren, an verschiedene andere Forscher geschickt. Noch 1947, also zwei Jahre nach Kriegsende, haben die Forscher, die von Mengele mit Blutproben versorgt worden waren, in einem Workshop zu den Abwehrfermenten in Tübingen über ihre damit gewonnenen Ergebnisse berichtet.

Obwohl eindeutig feststeht, dass die wichtigsten Entdeckungen von Abderhalden, der übrigens auch ein überzeugter Eugeniker war, ausschließlich auf Selbsttäuschung und Betrug beruhten, hat ihm dies weder zu Lebzeiten noch posthum geschadet. Interessant ist, dass dieser Forscher von 1922 bis 1935 eine Zeitschrift über Ethik herausgab. Er schrieb ein Lehrbuch der Biochemie, das zwischen 1906 und 1948 in 28 Auflagen erschien und in vier Sprachen übersetzt wurde. Er war von 1931 bis 1950, also bereits vor, aber auch noch 5 Jahre nach der Nazizeit, Präsident der Leopoldina, der bekanntesten deutschen wissenschaftlichen Akademie.

*„Es ist kein so sicherer Weg,
sich einen Namen zu machen,
als wenn man über Dinge schreibt,
die einen Anschein von Wichtigkeit haben,
die sich aber nicht leicht ein
vernünftiger Mensch die Zeit nimmt,
zu untersuchen“*

1.3.2 Beispiel 2: Der so genannte Chlorpropamid-Alkohol-Flush-Test

Mitte der 70er-Jahre tauchte ein etwas merkwürdiger Test in der Diabetologie auf, der so genannte Chlorpropamid-Alkohol-Flush-Test. David Pyke aus London hatte beobachtet, dass viele mit Chlorpropamid behandelte Diabetiker schon auf kleine Dosen von Alkohol, etwa ein Glas Sherry, eine spürbare Wär-

me und sichtbare Rötung im Gesicht entwickelten, die als Flush-Reaktion bezeichnet wurde. Immer wieder schien sich zu bestätigen, dass dieser Test fast nur bei Typ-2-Diabetikern positiv ausfällt. Dabei sollte er mit einem Diabetes mit guter Prognose bezüglich der Entwicklung von Spätkomplikationen assoziiert sein. Bei vielen Kollegen, unsere eigene Arbeitsgruppe eingeschlossen, fand sich immer wieder eine Bestätigung der Beobachtungen von Pyke. Im September 1979 hat Pyke die Claude-Bernard-Medaille der Europäischen Diabetes-Gesellschaft erhalten und in Wien zu diesem Thema einen geradezu brillanten Vortrag gehalten, der auch in einer der renommiertesten diabetologischen Zeitschriften abgedruckt wurde [3]. Dieser rhetorisch hervorragende Vortrag wurde mit lang anhaltenden stehenden Ovationen bedacht. Pyke war es gelungen, über die Flush-Reaktion, die über Enkephaline mediiert sein sollte, eine große Brücke zu Claude Bernard zu schlagen und hieraus bestechende Hypothesen zur Pathogenese des Typ-2-Diabetes abzuleiten. Alles war sehr überzeugend und schien gut belegt zu sein.

Wenig später kam bei uns in Göttingen aber trotzdem Zweifel auf, und wir fingen an, den Test doppelblind durchzuführen. Wichtig war dabei auch, dass der Untersucher nicht wusste, um welchen Typ des Diabetes es sich bei den Probanden handelte. Plötzlich wurden keinerlei Assoziationen mehr beobachtet [4]. Ich entsinne mich noch deutlich, wie wir uns damals immer wieder fragten, ob wir es überhaupt wagen könnten, derartige Ergebnisse zu publizieren angesichts der inzwischen erschienenen immensen Primär- und Sekundärliteratur zu diesem Thema. Unvergessen ist auch, welcher Unwille, ja welche geradezu ärgerlichen Reaktionen ausgelöst wurden, als wir zum ersten Mal im klinikinternen Kolloquium über die negativen Ergebnisse berichteten. Noch ganz unter dem Eindruck des erst wenige Wochen alten epochalen Festvortrages aus Wien wur-

de uns eine unangemessene jugendliche Lust am Widerspruch unterstellt. Inzwischen spricht niemand mehr vom Flush-Test, der allenfalls noch als kurioses Ereignis in der Geschichte der Diabetologie bekannt ist.

„Die gefährlichsten Unwahrheiten sind Wahrheiten, mäÙig entstellt“

1.3.3 Beispiel 3: Die retrolentale Fibroplasie

Anfang der 40er-Jahre stellte man zunächst in England, später auch in mehreren anderen Ländern eine bis dahin unbekannte Augenerkrankung bei Säuglingen fest, die meist in den ersten zwei bis drei Monaten auftrat und bevorzugt frühgeborene Kinder betraf. Veränderungen der BlutgefäÙe der Retina mit einer Bildung von fibrösem Gewebe hinter der Augenlinse führten bei den betroffenen Kindern zu einer Erblindung. Diese neue Erkrankung nahm schnell an Häufigkeit zu. Sie verschwand fast schlagartig wieder im Jahre 1955, nachdem das Konzept der Behandlung von Frühgeborenen geändert wurde [5]. Wir wissen heute, dass es sich um eine iatrogene Krankheit handelte, die durch eine übermäßige Sauerstoffzufuhr in den Brutkästen entstand. Zu dem damaligen Zeitpunkt wurde es erstmals möglich, Inkubatoren mit geringen Kosten mit einer hohen und relativ konstanten Sauerstoffkonzentration zu insufflieren. Die früher häufig beobachtete, mit Zyanose einhergehende Hypoxie der Frühgeborenen konnte auf diese Weise beseitigt werden. In Ermangelung genauerer Messmethoden wurde die Sauerstoffgabe an der Hautfarbe der Frühgeborenen ausgerichtet. Da man annahm, dass ein Sauerstoffmangel auch bei noch nicht zyanotischen Frühgeborenen zu Schäden führen könnte, erfolgte schließlich 1945 die Empfehlung, alle Frühgeborenen prophylaktisch mit Sauerstoff zu behandeln.

Die retrolentale Fibroplasie hielt man zunächst für eine angeborene Störung, am ehesten als Folge einer pränatal erlittenen Schädigung, wobei eine pränatale Anoxie als denkbare Ursache die meisten Anhänger fand. Durch methodisch sehr angreifbare Fallkontrollstudien wurden auch andere Ursachen erforscht, wie vorgeburtliche Infektionen, Blutungen, Mehrlingsschwangerschaften und anderes. Unter möglichen postnatalen Ursachen wurden eine zu frühe Lichtexposition, eine zu hohe Zufuhr von wasserlöslichen Vitaminen oder Eisen, ein Vitamin E-Mangel oder eine vorübergehende Nebennierenunterfunktion angeschuldigt. Nachdem die ACTH-Gabe bei einigen Kindern zu einer Besserung einer beginnenden retrolentalen Fibroplasie geführt hatte, wurde diese Therapie in größerem Umfang durchgeführt, wobei sich aufgrund von Beobachtungsmängeln und falschem Studiendesign die Wirksamkeit dieser Therapie lange Zeit zu bestätigen schien.

An die Möglichkeit eines Zusammenhangs mit der Sauerstofftherapie wurde trotz des eindeutigen Zusammenhangs zwischen dem Auftreten der „Epidemie“ und dem Beginn dieser neuen Therapie nicht ernsthaft gedacht. Zu sehr war man von dem Segen des Sauerstoffs überzeugt. Nachdem ein solcher Verdacht jedoch aufgrund einiger kleiner Studien aufgekommen war, wurden immer noch alternative Erklärungsmöglichkeiten bevorzugt. So wurde z.B. behauptet, es sei nicht der Sauerstoff selbst, sondern eine zu schnelle Entwöhnung vom Sauerstoff, die zu der Störung führte. Gegen die Durchführung randomisierter Studien wurden sehr heftig ethische Argumente vorgetragen. Nachdem in einer solchen Studie auch in der Gruppe der Frühgeborenen, die nicht oder nur sehr zurückhaltend mit Sauerstoff behandelt wurden, einige Fälle mit retrolentaler Fibroplasie auftauchten, wurde bekannt, dass die Krankenschwestern in der Nacht heimlich Sauerstoff zugeführt hatten, um die

nach ihrer festen Überzeugung so segensreiche Therapie den Säuglingen nicht vorzuenthalten. Als dann aber schließlich ordnungsgemäß durchgeführte Studien abgeschlossen werden konnten, wurde klar, dass der Verzicht auf Sauerstoff zum Verschwinden der retrolentalen Fibroplasie führte. Überraschenderweise ergab sich darüber hinaus, dass auch die allgemeine Mortalität durch die Reduzierung von Sauerstoff günstig beeinflusst wurde.

Diese sehr suggestive, aber in ihrer Wirksamkeit nicht ernsthaft überprüfte Sauerstoffgabe bei Frühgeborenen hat über einen Zeitraum von 12 Jahren schätzungsweise 10 000 Kinder das Augenlicht gekostet [5]. Dies wäre vermeidbar gewesen, wenn sehr viel früher Zweifel an dieser Therapieform aufgetaucht wären, oder wenn zumindest der erste aufgekommene Zweifel nicht so konsequent verdrängt worden wäre, sondern sofort zu einer konsequenten Überprüfung der Hypothese geführt hätte.

„Nichts setzt dem Fortgang der Wissenschaft mehr Hindernis entgegen, als wenn man zu wissen glaubt, was man noch nicht weiß. In diesen Fehler fallen gewöhnlich die schwärmerischen Erfinder von Hypothesen“

1.4 Gründe für fehlenden oder unterdrückten Zweifel bei der Gewinnung von wissenschaftlichen Erkenntnissen

Wie ausgeführt, gehört der Zweifel als unverzichtbarer Bestandteil zur Wissenschaft. Trotzdem unterdrücken Wissenschaftler häufig einen eigentlich angebrachten Zweifel, womit sie sich eindeutig unwissenschaftlich verhalten. Wie ist dies zu erklären? Nicht selten führt der Zweifel zur Gefährdung einer Hypothese und damit – zumindest vermeintlich – zur möglichen Zerstörung einer wissenschaftlichen Hypothese, die ggf. Basis für

den Ruhm des Wissenschaftlers ist. Sehr deutlich wurde dies am Beispiel von Abderhalden. Ein Zweifel an der Existenz der Abwehrfermente, der, wie wir heute wissen, eindeutig zur Widerlegung der Hypothese geführt hätte, hätte der wissenschaftlichen Bedeutung von Abderhalden wesentlichen Schaden zugefügt. Er hat deshalb jeden Zweifel mit der gesamten Kraft seiner Autorität unterdrückt. Dies wurde mit einer solchen Konsequenz und Skrupellosigkeit durchgeführt, dass die Vernichtung der Forscherkarriere anderer, in diesem Falle Michaelis, in Kauf genommen wurde. Sehr sorgfältig wurde von Abderhalden sein Gebäude von Lügen oder Selbsttäuschungen nach allen Seiten hin abgesichert.

Als Ursachen für ein solches Unterdrücken von Zweifeln durch Wissenschaftler kommen allgemeine schlechte menschliche Eigenschaften wie Eitelkeit, Dominanzstreben, unbändige Erfolgssucht und andere infrage. Gelegentlich scheint es so, dass ein solches wissenschaftliches „Kartenhaus“ (vielleicht auch besser Lügengebäude) ein derartiges Ausmaß erreicht hat, dass es dem Begründer der Hypothesen nicht mehr möglich erscheint, sich ohne vollständigen Gesichtsverlust hiervon zu lösen. Dies führt dann zu den schon geschilderten Absicherungen der Hypothesen durch Unterdrückung von Zweifel mit allen Mitteln.

In der heutigen wissenschaftlichen Welt, ganz besonders im Umfeld der Medizin, ist der wichtigste Faktor bei der Unterdrückung von Zweifel jedoch das Geld. Direkte persönliche Zuwendungen sind sicher eher die Ausnahme, wenngleich großzügige Reiseunterstützungen oder unangemessene Vortragshonorare durchaus im Grenzbereich zur persönlichen Zuwendung liegen. Geld spielt aber auch unmittelbar bei der Unterstützung der Wissenschaft eine ganz erhebliche Rolle, und so wird leicht ein angemessener Zweifel unterdrückt, wenn dieser dazu führen könnte, dass wesentliche Quellen der wissen-

schaftlichen Unterstützung dadurch versiegen könnten. An dieser Stelle sollen keine Beispiele genannt werden, aber jeder kritische Beobachter könnte auf Anheb Beispiele nennen, bei denen ganz offensichtlich finanzielle Zuwendungen zu einer mangelnden Bereitschaft zum Zweifel geführt haben. In den allermeisten Fällen würden die jeweils betroffenen Nichtzweifler einen solchen Zusammenhang allerdings heftig bestreiten. In vielen Fällen wird auch in der Tat der Zweifel nicht mehr bewusst unterdrückt. Diverse psychologische Mechanismen helfen vielmehr, das eigene Gewissen zu entlasten und den mangelnden Zweifel durch allerlei Sekundärbegründungen ethisch aufzuwerten. Insgesamt empfiehlt es sich, mit besonderer Wachheit auf möglicherweise unterdrückte Zweifel zu achten, wenn direkte oder indirekte geldliche Zuwendungen durch diejenigen erfolgen, denen der Zweifel ungelegen käme. Geldzuwendungen, die direkt oder indirekt zur Unterdrückung von Zweifeln führen, behindern die Wissenschaft. Sie sind daher immer unethisch, auch dann, wenn sie formal keinen Gesetzen widersprechen.

Insgesamt ist es also leicht erklärbar, wenngleich nicht entschuldbar, dass bei vielen Wissenschaftlern bezüglich ihrer eigenen Hypothesen Zweifel unterdrückt werden. Die Hypothesen werden aber erst zur Wissenschaft, wenn sie auch interpersonell Anerkennung finden. Dies führt zu der viel schwierigeren Frage, warum auf dem Gebiet der Wissenschaft auch im Umgang mit den Hypothesen und vermeintlichen Erkenntnissen anderer ein derart verbreiteter Mangel an angemessenem Zweifel herrscht.

„Zweifel muss nichts weiteres sein als Wachsamkeit“

1.5 Mangelnder Zweifel beim Umgang mit wissenschaftlichen Erkenntnissen

Wir haben gesehen, wie gefährlich im Umfeld von Abderhalden die Äußerung von Zweifel war, wollte man nicht seine eigene wissenschaftliche Karriere aufs Spiel setzen. Nachdem die wissenschaftliche Karriere von Michaelis einen dramatischen Knick erlitten hatte, ist es nachvollziehbar, dass andere Wissenschaftler ihren Zweifel nicht mehr offen aussprachen. Der später sehr berühmte organische Chemiker Brockmann konnte in Abderhaldens Institut einige Ergebnisse nicht reproduzieren. Als er Abderhalden hierüber unterrichtete, fragte dieser ihn, warum er ein Experiment, das einmal gut funktioniert habe, überhaupt wiederhole. Brockmann hat hieran schnell erkannt, dass Abderhalden ein Betrüger war, denn ein Experiment nie zu reproduzieren oder gar erfolgreiche Kontrollversuche zu verdrängen, ist nicht Wissenschaft, sondern Pseudowissenschaft oder Betrug. Er hat das Institut von Abderhalden sofort verlassen, leider allerdings ohne seine Zweifel öffentlich kundzutun. Auch dies ist, wenngleich menschlich verständlich, hochgradig unwissenschaftlich und nur durch ein extrem autoritäres Umfeld zu erklären. Auch aus anderen Gründen mag es gelegentlich opportun erscheinen, einen aufgekommenen Zweifel an der Hypothese eines anderen nicht zu äußern. Hier können die gleichen Mechanismen eine Rolle spielen, die oben bereits im Zusammenhang mit der Aufstellung von Hypothesen genannt wurden, etwa Karrierestreben oder finanzielle Abhängigkeit.

Viel häufiger sind jedoch Fälle, in denen ein Zweifel gar nicht erst aufkommt, wenngleich Gründe für einen Zweifel eigentlich

auf der Hand liegen. Wenn bestimmte Hypothesen erst einmal ein hohes Maß an Anerkennung gefunden haben – und insbesondere wenn sie von bedeutenden und charismatischen Menschen vertreten werden –, kann dies leicht dazu führen, dass sich ein Beobachter selbst Fesseln auferlegt, aufkommende Zweifel nicht wahrhaben will und es einfach nicht für möglich hält, dass bestimmte Aussagen oder Hypothesen nicht stimmen könnten. Nachdem 20 Jahre lang die wissenschaftliche Welt von Abwehrfermenten redet und eine Unzahl von Publikationen hierzu erschienen sind, ist es leicht verständlich, dass dem einzelnen Forscher der Mut fehlt, diesbezüglich Zweifel anzumelden. Wenn Tausende von Zuhörern nach einem glänzenden Vortrag über den Chlorpropamid-Alkohol-Flush-Test stehend applaudieren, dann wird so leicht kein einzelner Forscher den Mut für den Gedanken aufbringen, dass die gesamten diesbezüglichen Hypothesen schlichtweg vorgetäuscht sind. Wenn die medizinische Welt über viele Jahre die hochdosierte Sauerstoffzufuhr bei Frühgeborenen praktiziert, erfordert der Gedanke, dies könne zu schweren iatrogenen Schäden führen, viel Mut.

Die feste Überzeugung von der Wirksamkeit einer Therapie auf der Basis einer hohen Plausibilität kann besonders leicht zur Unterdrückung von Zweifeln bezüglich dieser Wirksamkeit führen und damit den Erkenntnisgewinn behindern. Ein Vorurteil wird umso stabiler, je mehr es zu den gängigen Vorstellungen der Medizin passt und je mehr die Falsifizierung der Überzeugung zu Veränderungen des ärztlichen Handelns führen müsste, insbesondere dann, wenn diese Veränderungen nach dem allgemeinen Vorurteil zu Gefährdungen führen. In solchen Situationen greifen die geschilderten Mechanismen der Täuschung und Selbsttäuschung, die zur immer wiederkehrenden scheinbaren Bestätigung der Hypothese führen und damit einen Zweifel nicht ernsthaft aufkommen lassen.

Die „Immunisierung“ gegenüber Zweifeln geschieht über verschiedene, zum großen Teil unbewusste Mechanismen. Auf vielen Ebenen der uns umgebenden Welt, in Politik, Werbung und Ideologien unterschiedlichster Art findet eine Veränderung statt: Aussagen werden immer weniger nach ihrer inhaltlichen Richtigkeit beurteilt werden und immer stärker nach ihrer Brauchbarkeit für bestimmte Ziele. Eine solche Denkweise ist bei vielen Menschen zur absoluten und nicht mehr reflektierten Selbstverständlichkeit geworden. Zweifel an der Richtigkeit einer Aussage können in einem derartigen Umfeld fast immer nur störend sein. Diese Denkweise findet immer weiteren Eingang auch in die Medizin. Aussagen, die dem Ziel der sozialen Absicherung, der Karriereförderung und damit der Einkommensmaximierung dienen, werden unabhängig vom Wahrheitsgehalt stabil gehalten. Solange man sich bezüglich dieser Aussagen vor Zweifeln schützt, muss nicht einmal eine bewusste Verdrehung von Tatsachen in Kauf genommen werden.

Ein anderer Schutzmechanismus vor Zweifeln ist psychologisch durchaus nachvollziehbar. Der Ärger, der den Zweifler trifft, resultiert nämlich oft auch aus einer Projektion desjenigen, der selber kurz zweifelte, sich diesen Zweifel aber erfolgreich versagte. Täglich mit einer Sache umzugehen und diese vertreten zu müssen, obwohl sie erkennbar nicht funktioniert, führt zu einer kognitiven Belastung, die kein Mensch lange aushält. Man passt sich in einer Art Schutzreflex unwillkürlich an und lässt nur Erkenntnisse in das Bewusstsein, die mit der Lebenssituation konform sind. Wenn man schon, aus welchen Gründen auch immer, ein Medikament ohne nachgewiesene Wirksamkeit verschreibt, dann möchte man sich nicht mit dem bewussten Gefühl belasten, etwas Unärztliches oder zumindest etwas Unwirtschaftliches getan zu haben. Man denkt über die Frage der Wirksamkeit lieber nicht nach

oder nimmt nur solche Informationen auf, die mit einer Wirksamkeit des Medikamentes in Einklang wären. Insofern ist Evidenzbasierte Medizin für Menschen, die die Kultur des Zweifels nicht beherrschen, unbequem oder gar bedrohlich.

Schädlich für die Kultur des Zweifels und die damit verbundene wissenschaftliche Denkweise in der Medizin ist ferner die modische Sucht nach „Konsensuskonferenzen“ bzw. „Konsensus-Statements“. Der Soziologe Hondrich hat vor einiger Zeit in einem Spiegel-Essay über die potenzielle Wissenschaftsfeindlichkeit gesellschaftlicher Konsense geschrieben. Er führte aus, wie stabil ein so genannter Wertekonsens sei, für den der Wahrheitssucher, also der zweifelnde Wissenschaftler, der schlimmste Feind sei. Dies gilt auch für die Medizin, in der bestimmte Konsense durchaus für den wissenschaftlichen Fortschritt hinderlich sein können, wenn Äußerungen, die einem solchen Konsens zuwiderlaufen, mit einem Bann belegt werden. Dem Zweifler droht gewissermaßen der Ausschluss aus der wissenschaftlichen Gemeinschaft. Es ist viel bequemer, im Konsens die Unwahrheit zu sagen als im Streit die Wahrheit zu suchen.

*„Man spricht viel von Aufklärung und
wünscht mehr Licht.*

*Mein Gott, was hilft aber alles Licht,
wenn die Leute entweder keine Augen haben
oder die, die sie haben,
vorsätzlich verschließen“*

1.6 Des Kaisers neue Kleider

Falsche Hypothesen, seien sie durch Selbsttäuschung oder durch bewusste Verdrängung von Zweifel entstanden, gewinnen also über verschiedene Mechanismen eine Art der Selbsterhaltung oder der inneren Stabilisierung, wenn sie nur eine gewisse Größe erreicht haben und wenn als Folge davon

kaum jemand es für möglich hält, dass alles nur auf Irrtümern oder einem Lügengebäude beruht. Es handelt sich hierbei um das Phänomen von „des Kaisers neuen Kleidern“, wie es in dem gleichnamigen Märchen von Andersen beschrieben wurde. Wenn auch die Annahme über die nur für dumme Menschen unsichtbaren Kleider noch so absurd war, niemand der Erwachsenen brachte den Mut auf, einen Zweifel zu äußern. Unterstützt wurde dies durch die Behauptung der kaiserlichen Vasallen, dass derjenige, der die Kleider nicht sieht, zu erkennen gibt, dass er dumm sei. Für die Bevölkerung war es einfacher, mit der gesamten übrigen Masse zu stimmen, als den eigenen auf der Hand liegenden Zweifel zu äußern und damit das Risiko einzugehen, für dumm gehalten zu werden. Erst die naive Frage des Kindes, das den eigenen Zweifel ganz unschuldig aussprach, brachte das kollektive Kartenhaus zum Zusammenbruch.

Das gleiche Verhaltensmuster gilt für die Wissenschaft: Viele Zweifel werden nur deshalb nicht geäußert, weil der Zweifler angesichts der großen Masse der Andersdenkenden Angst hat, für dumm gehalten zu werden. In der Tat ist der Umgang mit dem Zweifler in der wissenschaftlichen Welt häufig wenig zimperlich. So werden ihm nicht nur Dummheit, sondern destruktive Kritiksucht, modisches Infragestellen oder sogar unethische Motive unterstellt. Derartige Argumentationsmuster, die damit indirekt zur Unterdrückung von Zweifeln beitragen, sind jedoch nicht mit dem Geist der aufrichtigen Wissenschaft vereinbar. Nur dort, wo sich der wissenschaftliche Zweifel frei entfalten kann, kann die Wissenschaft blühen und gedeihen.

*„Man soll alle Dinge im Leben
mindestens einmal bezweifeln“*

1.7 Schlussfolgerung

Ausgehend von der Erkenntnis, dass der wissenschaftliche Fortschritt zum großen Teil auf der Spirale „Hypothese – Zweifel – Falsifikation – neue Hypothese“ beruht, kommt dem Zweifel eine zentrale Bedeutung im Prozess des Erkenntnisgewinns zu. Der Wissenschaftler, der – aus welchen Gründen auch immer – im Zusammenhang mit seinen eigenen Hypothesen den Zweifel unterdrückt, handelt unwissenschaftlich. Aber auch im Umgang mit den Hypothesen anderer ist permanenter Zweifel angezeigt und, wenn angemessen formuliert, immer ethisch hochstehend. Unterdrückung von Zweifel ist dagegen unethisch und verwerflich. Viele unterschiedliche Mechanismen, die in der Person des verhinderten Zweiflers selbst oder in gesellschaftlichen Strukturen liegen können, führen zu unangemessener Unterdrückung von Zweifeln. Solche Mechanismen aufzudecken und ihnen entgegenzuwirken, ist eine wichtige Aufgabe aller Wissenschaftler. Nur in einem Umfeld, in dem der Zweifel gedeihen kann, gedeiht auch die Wissenschaft.

Dies gilt insbesondere auch für die Wissenschaft in der Medizin. Auch hier fördert nur der Zweifel die Suche nach Evidenzen. Insofern ist der Zweifel auch ein wesentlicher Bestandteil der so genannten evidence based medicine, die ihrerseits ja wissenschaftliche Belege zur Grundlage des ärztlichen Handelns macht. Daraus ergibt sich folgerichtig, dass ein wesentliches Merkmal der Evidenzbasierten Medizin darin liegt, dass sie sich in all ihren Aussagen immer wieder selbst in Zweifel zieht und nach besseren und neueren Belegen für ärztliches Handeln sucht.

Literatur

- [1] Popper KR (1982) Logik der Forschung. Mohr, Tübingen
- [2] Deichmann U, Müller-Hill B, The fraud of Abderhalden's enzymes. Nature. (1998), 393, 109–111
- [3] Pyke DA, Diabetes: the genetic connections. Diabetologia. (1979), 17, 333–343
- [4] Köbberling J et al., The chlorpropamide alcohol flush. Lack of specificity for familial non-insulin dependent diabetes. Diabetologia. (1980), 19, 359–363
- [5] Jacobson RM, Feinstein AR, Oxygen as a cause of blindness in premature infants: „autopsy“ of a decade of errors in clinical epidemiologic research. J Clin Epidemiol. (1992), 45, 1265–1287

2 Theorie, Geschichte und Ethik der Evidenzbasierten Medizin (EbM)

Heiner Raspe

2.1 Zugänge zur EbM

Man kann sich dem Verständnis der EbM auf ganz unterschiedlichen Wegen nähern, über

- ▲ die Praxis der Medizin,
- ▲ die medizinische Forschung,
- ▲ die Theorie der Medizin,
- ▲ die Geschichte der Medizin,
- ▲ aktuelle Selbstverpflichtungen der ärztlichen Profession und
- ▲ ethische Überlegungen.

Im Folgenden soll jeder dieser Zugänge eröffnet und ein Stück weit entwickelt werden. Dabei geht es allein um die evidenzbasierte klinische Medizin. Zur evidenzbasierten gesundheitlichen Versorgung (Evidence-based health care, EBHC) sei auf das Buch von Muir Gray (2. Aufl. 2001, [1]) verwiesen.

2.2 Evidenzbasierte Praxis der Medizin

EbM ist zuerst und vor allem klinische Medizin, d.h. medizinische (ärztliche, aber auch pflegerische, physiotherapeutische, psychotherapeutische etc.) Praxis in direktem Kontakt mit und direkter Verantwortung für individuelle Patienten – gleichgültig, ob in einem Krankenhaus oder einer Praxis. EbM hat keine besonderen Patienten, keine anderen Medikamente, Röntgeneräte, Krankenstationen, Praxisräume als Medizin sonst auch. Sie kennt keine anderen Ziele: Prävention, Gesundheitsförderung, Lebenserhaltung, Heilung, Linderung, Trost, Verhütung des Fortschreitens von Krankheiten, Palliation, Pfl-

ge, bedingt gelingendes Gesund-Sein, Ermöglichung eines guten Todes (nach [2–4]). Sie ist keine besondere Therapierichtung!

Entsprechend haben der Internist und Nephrologe David Sackett et al. die EbM 1996 [5] definiert als

„the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external evidence from systematic research.“

Entscheidungen über die Behandlung individueller Patienten trifft jeder praktisch tätige Arzt jeden Tag, selbstverständlich auch unter Berücksichtigung der klinischen Situation und der Präferenzen und Verhaltensweisen der Patienten. Was die EbM spezifisch hinzufügt, verdeutlicht das 2002 von B. Haynes et al. [6] aktualisierte Modell evidenzbasierter klinischer Entscheidungen unter dem Untertitel: „Evidence does not make decisions, people do.“

Gemeint ist das, was in dem Modell „research evidence“ und in der o.g. Definition „best available external evidence from systematic research“ heißt. Entscheidungen gründen also – zumindest auch – auf „evidence“, und zwar auf „external evidence“.

Das Wort „external“ erlaubt verschiedene Interpretationen. Es drückt auch aus, dass die Research-Evidence von außen zur klinischen Situation hinzutritt, sozusagen importiert werden muss. Sie ist, anders gesagt, nicht aus der Situation heraus generierbar. Dies wird



Abb. 2.1: Ein Modell evidenzbasierter klinischer Entscheidungen

besonders deutlich bei der Beurteilung langfristiger präventiver Interventionen. Ob z.B. Statine die Inzidenz von Herzinfarkten reduzieren, ist der alltäglichen (Bacon nannte sie „incondita/roh“) und auch der reflektierten klinischen Erfahrung Einzelner unter keinen Umständen zugänglich. Wir haben kein Organ für ausbleibende seltene Ereignisse. Zur Beantwortung solcher Fragen bedarf es mehrjähriger kontrollierter Studien.

2.3 EbM und klinische Forschung

Praxisrelevante Forschung wird in der Medizin seit langem mit dem Adjektiv „klinisch“ belegt. In Anlehnung an Spitzer [7] kann man klinische Forschung definieren als das Studium der Determinanten und Folgen klinischer Urteile, Entscheidungen und Aktionen. Sie antwortet auf die handlungsleitenden ärztlichen Fragen: Was liegt vor? Was muss ich (bei Handlungsverzicht) befürch-

ten? Was soll ich tun und darf ich dann hoffen? Worauf muss ich achten? Jede Frage lässt sich – mit Unschärfen – einem klinischen Handlungsbegriff (Diagnose – Prognose – Prävention/Therapie/Rehabilitation/Palliation – Verlaufskontrolle) und einem bestimmten Studientyp zuordnen (Querschnittstudie – Kohortenstudie – randomisierte klinische Studie/RCT – Outcomes-Beobachtung). Hier zeigt sich ein erstes Mal, dass EbM mehr braucht als die Ergebnisse confirmatorisch angelegter RCTs. In jüngster Zeit tritt ein RCT-Typ in den Vordergrund, der schon 1967 von Schwartz und Lellouch [8] mit dem Zusatz „pragmatic attitude“ charakterisiert wurde [9]. Heute spricht man von Practical clinical trials [10] oder Effectiveness-Studies [11]. Treweek et al. [12] haben ein Trial protocol tool entwickelt, um die Protokollentwicklung von Pragmatic randomized controlled trials zu unterstützen.

EbM wird auch nicht auf Ergebnisse der ätiologischen und pathogenetischen biome-

dizinischen Forschung verzichten können; denn oft gibt die Beantwortung der Frage nach dem, was „hinter dem klinischen Bild steckt“, diagnostisch, prognostisch und therapeutisch wichtige Hinweise. Und schließlich muss ein zunehmend wichtiger werdender Forschungstyp erwähnt werden, die qualitative Studie. Ihre Ergebnisse machen verständlich, warum Patienten Ärzte aufsuchen oder eben auch nicht, warum sie sich so (aus ärztlicher Sicht oft irrational) verhalten, wie sie sich verhalten, warum auch Ärzte „unfolgsam“ sind und z.B. Leitlinien nicht folgen mögen (s. Kap. 35). Die nicht aus RCTs abzuleitende Evidenz ist wichtig, vor allem im Übergang von der Beweiswürdigung (Quality of evidence) zur mehr oder minder starken Handlungsempfehlung (Strength of recommendation) [13; 14].

Die Deutsche Forschungsgemeinschaft unterschied in ihrer Denkschrift 1999 [15] drei Aspekte der klinischen Forschung: die Grundlagen-, die Krankheits- und die Patienten-orientierte Forschung. Von diesen ist die letzte für die EbM sicher am bedeutendsten. Sie erfordert nach einer Definition des National Institute of Health (NIH) einen direkten Kontakt zwischen Arzt und Patient, „working together, both being warm and alive“ [16]. Dort, wo EbM die Folgen klinischen Handelns (und damit auch einen zentralen Teil seiner Determinanten) untersucht, fixiert sie einen besonderen Typ ärztlicher Erfahrung, die wissenschaftlich kontrollierte Erfahrung (in Abgrenzung zur rohen und zur reflektierten Erfahrung). Eine Bemerkung Paul Martinis [17] verdeutlicht die Bedeutung einer wissenschaftlichen Kontrolle:

„Im körperlichen wie im seelischen Bereich ist der ärztliche Erfolg nicht damit bewiesen, dass ein Kranker gesund geworden ist. Wenn ein Kranker nicht gesund wird, dann hat der Arzt immer Grund zu sagen, dass er einen Misserfolg gehabt habe. Das heißt noch nicht, dass

er die Ursache der Nicht-Heilung gewesen sei. Wenn ein Kranker aber gesund wird, dann hat ein Arzt fürs Erste ebenso wenig das Recht, seinem Eingreifen das Verdienst für die Heilung zuzuschreiben. Das kann er nur dann tun, wenn er in seiner Versuchsanordnung die Voraussetzung für einen folgerichtigen Schluss geschaffen hat.“

Insofern ist Evidenz mehr als das, was in Veröffentlichungen an Daten in Tabellen und Grafiken festgehalten ist. Liest man die „Evidenz“ der EbM als „evidence“, d.h. als Indiz bzw. Beweismittel, dann kommt es am Ende auf eine kritische Beweiswürdigung an. Dabei spielt die Versuchsanordnung eine zentrale Rolle. Aber auch die Auswahl der Versuchspersonen und des Settings, die Stichprobengröße, die Qualität der Versuchsdurchführung, die Wahl und Messung der Ergebnisindikatoren und die Dauer der Nachbeobachtung sind wichtige Gegenstände des Critical Appraisal. Veröffentlichte Daten können nur dann zu überzeugender Evidenz (Convincing, persuasive evidence) werden, wenn sie einer kritischen Diskussion standgehalten haben.

Die Versuchsanordnung ist entscheidend für die Frage, ob eine untersuchte Intervention wirksam ist, ob die beobachteten Effekte also ganz oder teilweise auf sie als Ursache zurückgeführt werden können. Hier kann die Antwort im Prinzip nur „ja“, „nein“ oder „unsicher“ lauten.

Eine zweite Frage ist, ob eine wirksame Intervention Nutzen gestiftet hat und weiter Nutzen stiften kann. Dazu sind Qualitäten, Quantitäten und Unsicherheiten der auf sie zurückzuführenden Effekte zu bewerten (entsprechend klinisch relevante Endpunkte vs. Surrogatparameter; Absolute und Relative Risiko-Reduktion, Number needed to treat; Konfidenzintervalle). In der Bilanz kommt es dann auf den Nettonutzen an, d.h. auf das,

was vom Nutzen übrig bleibt, wenn man alle der Intervention zuzuschreibenden Belastungen und Risiken abzieht. Schließlich ist zu berücksichtigen, welche Rolle veränderte Settings und Kontexte spielen. Lässt sich der Nettonutzen auch in meiner Praxis, in meinem Krankenhaus, für meine Patienten realisieren oder gibt es Gründe, die gegen eine unmittelbare Übertragbarkeit sprechen? Sie würden dazu führen müssen, die Handlungsempfehlungen den lokalen Verhältnissen anzupassen [18].

Daten und Evidenz werden primär in so genannten Originalien veröffentlicht. Was nicht veröffentlicht ist, ist für die EbM nicht existent. Die Ergebnisse von Primärstudien können in systematischen Übersichtsarbeiten mit und ohne Meta-Analysen verdichtet werden (Sekundärforschung). Kaum jemals reicht eine einzelne Studie aus, um Entscheidungen (vor allem von der Routine abweichende Entscheidungen) zu begründen. Allerdings hütet sich die EbM davor, die Abwesenheit von ausreichender und belastbarer Evidenz für Evidenz für die Abwesenheit von Wirksamkeit und Nutzen zu halten. Andererseits ist nicht alles, was veröffentlicht ist, ist für sie von Bedeutung. In der Regel werden zur Auswahl von Studien Qualitätsschwellen definiert, die nicht unterschritten werden dürfen. Die EbM orientiert sich an den besten erreichbaren Studien, der besten erreichbaren Evidenz, nicht am besten Vorstellbaren.

Sie tut dies nicht kritiklos: einerseits hat sie zahlreiche Techniken entwickelt, um erreichbare Evidenz auch tatsächlich zu erreichen. Eine einfache Suche über PubMed genügt nicht. Zweitens verfügt sie über Instrumente, die es erlauben, die Qualität von Studien und Veröffentlichungen strukturiert zu bewerten, und drittens gibt es (noch nicht vollständig befriedigende) Verfahren, um Hinweise auf unterdrückte Publikationen (Publication-Bias) zu gewinnen. Schließlich ist ihr bewusst geworden, dass zur Formulie-

rung realitätsgerechter Empfehlungen auch Studien geringer Evidenzgrade notwendig sind – im Bereich der Therapie etwa Studien aus Behandlungsfall-Registern, Qualitätssicherungs-Programmen, Anwendungsbeobachtungen oder Ergebnisse der Outcomes-Forschung.

2.4 EbM und die Theorie der Medizin

Handlungssysteme, deren Ziel es ist, den natürlichen Verlauf individueller menschlicher Entwicklungen in direktem Kontakt mit ihnen in eine als positiv bewertete Richtung zu verändern, gehören, wenn sie wissenschaftlich fundiert sind, zu den so genannten Praxis- oder Handlungswissenschaften. Sie sind zugleich Erfahrungswissenschaften und, wie sich zeigen wird, auch Normwissenschaften.

Ein zentrales Merkmal ihrer Wissenschaftlichkeit besteht darin, dass sie sich der unsicheren Folgen ihrer Handlungen immer wieder methodisch kontrolliert versichern, entweder formativ (unter regelmäßiger Beobachtung von Zwischenergebnissen) oder summativ (d.h. zusammenfassend nach Abschluss einer Intervention) [19]. Beide Typen der Evaluation sind aus wissenschaftstheoretischer Sicht konsequenzialistisch (orientieren sich an Handlungsfolgen), empiristisch (orientieren sich an messbaren Folgen) und probabilistisch (beschränken sich auf Wahrscheinlichkeitsaussagen).

In der klinischen Praxis steht die formative Evaluation (z.B.: „Wie reagiert mein Patient auf das neue Medikament?“), in der klinischen Forschung die summative Evaluation im Vordergrund.

Wollen Handlungswissenschaften in diesem Sinne wissenschaftlich sein und bleiben, dann lässt sich für sie eine Verpflichtung zu beständiger evaluativer Forschung ableiten. Patienten und Ärzte sind darauf angewiesen, Chancen und Risiken medizinischer Inter-

ventionen ex post beurteilen und ex ante abschätzen zu können.

Eine Handlungswissenschaft eigenen Rechts [20] ist die Medizin. Für sie gilt Grundsatz 6 der Helsinki-Deklaration:

„Oberstes Ziel der medizinischen Forschung am Menschen muss es sein, prophylaktische, diagnostische und therapeutische Verfahren sowie das Verständnis der Ätiologie und Pathogenese der Krankheiten zu verbessern. Selbst die am besten erprobten Methoden müssen fortwährend durch Forschung auf ihre Effektivität, Effizienz, Verfügbarkeit und Qualität überprüft werden.“

Kann man sich eine Medizin vorstellen, die auf solche prüfende und bewertende Forschung verzichtet? B. Naunyn [21] konnte es nicht; er erweiterte nach dem Ende seiner klinischen Tätigkeit (1905) seinen klassischen Aphorismus aus der Zeit seiner Habilitation:

„Für mich ist es kein Zweifel, daß das Wort: ‚Die Medizin wird eine Wissenschaft sein, oder sie wird nicht sein‘ auch für die Therapie gelten muß und gilt. Die Heilkunde wird eine Wissenschaft sein, oder sie wird nicht sein! Mir ist sonnenklar, dass dort, wo die Wissenschaft aufhört, nicht die Kunst anfängt, sondern rohe Empirie und das Handwerk.“

Andere Handlungswissenschaften sind die klinische Psychologie, die Pädagogik und die Rechtsprechung und wahrscheinlich auch die Seelsorge – mit je eigenen Situationen, Personen, Aufgaben und Zielen, Mitteln, Begründungen und Regulativen [22]. Die intendierten Veränderungen sind nicht ziellos und haben in jedem Fall einen persönlichen, zwischenmenschlichen und sozialen Wert. Auch und vor allem in der Medizin gilt, dass

jede klinische Feststellung – mehr oder weniger offensichtlich – eine Wertstellung enthält. Fieber über 40°C ist so gut wie immer negativ, etwas das gesenkt werden muss. Stellt man eine Senkung auf 38°C fest, ist dies zugleich eine Besserung. Ähnlich ist ein Tumor im Stadium T1N0M0 nicht nur etwas anderes als ein Tumor im Stadium T1N1M1, sondern er ist unmittelbar auch prognostisch günstiger. Und noch besser wäre es, es gäbe überhaupt keinen Tumor.

Es kann gut sein, dass ein Mensch gleichzeitig die Hilfe mehrerer Handlungswissenschaften in Anspruch nehmen muss. Dies macht deutlich, dass jede einzelne immer auch Reduktions- und Aspektwissenschaft ist; jede widmet sich, bei manchen Überschneidungen, doch „nur“ einer oder wenigen Seiten menschlicher Existenz (selbst wenn jede gerne mit dem Anspruch der „Ganzheitlichkeit“ daherkommt). Eine weitere Gemeinsamkeit ist darin zu sehen, dass alle einem Handlungszwang unterliegen. In der Medizin spricht man vom „therapeutischen Imperativ“ [23; 24]. Er ist hier so stark, dass jeder Handlungsverzicht einer besonderen Bezeichnung und Begründung bedarf („Oudenotheapie“, „gezieltes Zuwarten“, „to be masterly in-active“).

In ein solches Verständnis von Medizin als Praxiswissenschaft fügen sich Konzept und Methodologie der EbM nahtlos ein: das in der klinischen Situation unvermeidliche ärztliche Handeln begründet und legitimiert sich auch – aber keineswegs allein (s. Abb. 2.1) – mit bzw. durch Evidenz aus empirischer Forschung. Deren Prototyp ist die kontrollierte klinische Studie. Als Methodenlehre solcher Forschung ist die klinische Epidemiologie entwickelt worden [25–27]. Sie ist historisch eine Entwicklung von Klinikern (wie dem Internisten Paul Martini mit seiner Methodenlehre der therapeutischen Untersuchung [28; 29]) für Kliniker.

Die folgende Übersicht (Tab. 2.1) führt zentrale klinische Fragen auf und ordnet ih-

Tab. 2.1: Handlungsleitende Fragen der klinischen Medizin und ihr wissenschaftliches Fundament

Frage	Wissenschaftlicher Zugang
Was liegt in diesem Fall – objektivierbar – vor?	Klini- und Psychometrie
Wie häufig ist dieses Krankheitsbild hier (und anderswo)?	Prävalenz- und Inzidenzstudien
Wie gehe ich diagnostisch vor? Welche Tests soll ich auswählen?	Diagnostische Studien
Kann dem Krankheitsbild eine bestimmte Diagnose (Klasse, Grad, Stadium ...) zugeordnet werden?	Nosographie und Nosologie, Assessment, Klassifikation
Und eine Prognose?	Prognostische Studien
Was steckt dahinter? Ätiologie? Pathogenese?	Biologische und psychologische Grundlagenforschung
Was sind Motive, Anliegen und Präferenzen des Patienten?	Qualitative (hermeneutische) Studien, „narratives“
Was ist das Behandlungsziel?	Menschenbilder, ethische und rechtliche Normen
Was ist präventiv, therapeutisch, rehabilitativ zu tun, was hat sich (nicht) bewährt? In Studien und im Alltag?	Interventionsstudien, Anwendungsbeobachtungen, Outcomes-Research
Worauf muss ich im Verlauf achten?	Verlaufsbeobachtungen, Kasuistiken
Welche Kosten sind mit all dem verbunden?	Kosten-, Kosten-Nutzen-Studien, Modellierungen

nen entsprechende Forschungsfelder und Studiendesigns zu.

Vergegenwärtigt man sich die zur Beantwortung der einzelnen Fragen notwendigen Forschungsansätze, dann erscheinen Medizin und klinische Forschung als sehr variantenreiche Unternehmungen. Die Spanne der wissenschaftlichen Zugänge reicht von Kasuistiken (z.B. im Bereich der Identifikation von Nebenwirkungen) und hermeneutischen Studien über kontrollierte Interventionsstudien bis hin zu hoch abstrakten gesundheitsökonomischen Modellierungen. Medizin lässt sich offensichtlich nicht über einen wissenschaftlichen Leisten schlagen.

Mit Blick auf die rechte Spalte fällt auf, in welchem Ausmaß die nur eine Zeile einnehmende biomedizinische Forschung dennoch das Selbstverständnis, das Ansehen, die Ausbildung, die Forschungsaktivitäten und Forschungsförderung von Ärzten dominiert. Dies geschieht nicht völlig zu Unrecht, er-

scheint aber doch unbalanciert. Innovationen in der Medizin verdanken sich in weiten Bereichen der biologischen Grundlagenforschung und ihren Ergebnissen. Allerdings muss in jedem Fall geprüft werden, ob eine Innovation auch einen Fortschritt bedeutet, also wirksam ist, Nutzen ohne wesentlichen Schaden stiftet und eine Behandlungslücke schließt. Und genau zu dieser Prüfung bedarf es grundsätzlich anderer wissenschaftlicher Ansätze und Methoden als sie die biologische Grundlagenforschung bereitstellt.

2.5 EbM in medizinhistorischer Sicht

Das Selbstverständnis der Medizin als einer sich an ihren Folgen selbst kontrollierenden Handlungswissenschaft ist jung, auch wenn sich kontrollierte klinische Studien bis in die Mitte des 18. Jahrhunderts zurückverfolgen

lassen. Ein Meilenstein war die Studie des schottischen Marinearztes James Lind, berichtet in seinem „Treatise of the Scurvy“ aus dem Jahr 1753 (cf. www.jameslindlibrary.org). Auf weitere britische Vorläufer geht U. Tröhler in seinem Buch „To improve the evidence of medicine. The 18th century British origins of a critical approach“ ein [30]. Und auch die randomisierte klinische Studie (RCT) erscheint als eine (spätere) britische Erfindung. Ihr Geburtsjahr wird allgemein mit 1948 angegeben [31–33]. Dabei sollte man kontinental-europäische Vorläuferstudien nicht übersehen, darunter auch die des deutschen Adolf Bingel aus dem Jahr 1918 [34]. Es handelte sich um eine vermutlich doppelblinde kontrollierte Studie zur Serumbehandlung der Diphtherie mit alternierender Zuteilung.

Dennoch sah sich die neuzeitliche Medizin über Jahrhunderte vor allem als eine Erklärungswissenschaft im Kontext dogmatischer Systeme, wie sie bis über die romantische Medizin hinaus wirksam waren und sind. „Im Wissen, nicht im Können liegt unsere Kraft“, soll es in der neuen Wiener Schule (J. Dietl; 1804–1878) geheißen haben. Der tiefe Graben zwischen (vermeintlichem) Wissen und (geringem) praktischem Vermögen wurde hingenommen. Dafür wurde der polemische Begriff „therapeutischer Nihilismus“ geprägt.

Sicher waren es vor allem die Grundlagen- und die Krankheits-orientierte Forschung, die das spätere Können vorbereiteten. Dies ohne Einschränkung anzuerkennen, darf aber nicht dazu führen, das innovative Potenzial der klinischen Forschung im oben definierten Sinn zu übersehen. Es kommt dort, z.B. in so genannten Proof of concept-Studien, zum Tragen, wo das pathophysiologische Wissen noch hypothetisch ist, Lücken aufweist oder auch in der Rückschau schlicht falsch war. Ein anschauliches Beispiel gibt das Medikament Sulfasalazin (eine Kombination eines damals verfügbaren

Sulfonamids und der Salicylsäure), das in Schweden Anfang der 1940er-Jahre unter einer bis heute nicht belegten Infektonshypothese gegen die rheumatoide Arthritis entwickelt wurde [35]. Anfangs als enttäuschend wirkungsschwach beurteilt, wurde es dann für die Behandlung entzündlicher Darmerkrankungen entdeckt und um 1980 ein zweites Mal in die Rheumatologie eingeführt. Die ersten RCTs überraschten mit vergleichsweise positiven Ergebnissen; heute zählt Sulfasalazin zu den häufig gebrauchten langfristigen wirksamen Antirheumatika, mit weiterhin nicht vollständig geklärtem Wirkmechanismus. Und auch bei den schon erwähnten Statinen ist unsere Überzeugung, dass sie präventiv wirken (auch bei Personen mit normalen Cholesterinspiegeln) größer als unser Verständnis ihrer Pharmakodynamik.

Damit können die klinische und die Versorgungsforschung im Kontext der EbM selbst zu Fortschritten führen. Sie prüfen die Wirksamkeit von Methoden und Programmen auch dann, wenn ihre Wirkmechanismen noch unbekannt, unsicher oder umstritten sind. Damit regen sie die Grundlagenforschung an und geben ihr bestimmte Richtungen; sie erlauben auch die Prüfung komplexer Interventionen (z.B. stationäre Rehabilitation, Mammographie-Screening), deren Einzelkomponenten nur schwer voneinander isoliert werden können und die in ihren vielfältigen Interaktionen kaum übersehbar sind.

In Deutschland war es das Verdienst Paul Martinis (1889–1964; [36]), als erster eine „Methodenlehre der therapeutischen Untersuchung“ (später der „therapeutisch-klinischen Forschung“) entwickelt und 1932 veröffentlicht zu haben – 5 Jahre vor Hills „Principles of Medical Statistics“ [37]. Die Methodenlehre erlebte bis 1968 vier Auflagen, um dann in Deutschland als Buch und Programm in Vergessenheit zu geraten. Stoll et al. [28] werten seine Arbeit als eine „deutsche Vorgeschichte der Evidenzbasierten Medizin“.

Die internationale Geschichte der EbM, besonders des Begriffs, ist nicht sicher zu klären (D. Sackett pers. Mitteilung 1998). Er taucht als solcher in den 1980er-Jahren in der nordamerikanischen Literatur auf. Als Wiege der EbM kann die McMaster-Universität in Hamilton/Ontario in Kanada (Department of Clinical Epidemiology and Biostatistics) gelten, wo D. Sackett, B. Haynes und P. Tugwell forschten, klinisch tätig waren und lehrten. Andere frühe Zentren befanden sich in New Haven [38], Seattle [39] und Montreal [40].

Somit erweist sich die EbM als eine Entwicklung aus der klinischen Praxis und für die klinische Praxis auf dem Boden der Methodenlehre der klinischen Epidemiologie als „basic science for clinical medicine“ oder auch als „science for the art of medicine“ [41]. In diesem Sinne hatte D. Sackett einen klinischen Epidemiologen schon 1969 definiert als: „An individual with extensive training and experience in clinical medicine who, after receiving appropriate training in epidemiology and biostatistics, continues to provide direct patient care in his subsequent career.“

Bis die EbM Deutschland erreichte, vergingen rund 15 Jahre. Der erste, der auf diese nordamerikanische und dann bald auch britische Entwicklung bei uns aufmerksam machte, war wahrscheinlich der Sozialmediziner David Klemperer mit einem 1995 veröffentlichten Aufsatz über „Qualität und Qualitätskontrolle in der Medizin“ (dort im Abschnitt 4 über Klinische Entscheidungsfindung – Clinical decision making). Andere frühe Rezipienten kamen aus der Klinik [42; 43], der ärztlichen Selbstverwaltung [42] und wieder der Sozialmedizin [44]. Die klinisch tätigen Ärzte waren und sind in ihrer Mehrzahl eher auf der Seite der Kritiker zu finden. EbM hat in Deutschland bisher vor allem Karriere in Form der evidenzbasierten Versorgung gemacht. Schon 1999 fand sie Eingang in das Sozialgesetzbuch V, im inzwischen aufgehobenen § 137e („evidenzbasierte Leitlinien“), nachdem der Bundesausschuss der Ärzte und

Krankenkassen sich bereits im Oktober 1997 in seinen so genannten NUB-Richtlinien auf den Gebrauch von „Evidenzstufen“ festgelegt hatte. Die Bedeutung des Konzepts und der Methoden der EbM ist sozialrechtlich bis heute fortlaufend gestärkt worden. Eine milde Skepsis hat bisher nur das Bundesverfassungsgericht erkennen lassen.

Damit erschien EbM in Deutschland im Gewand einer „Revolution von oben“. Dieser Geburtsfehler ist bis heute nicht behoben. EbM konnte unter diesen Bedingungen kaum als das wahrgenommen werden, was sie historisch gesehen ist: eine intraprofessionelle Entwicklung der Selbstvergewisserung und Selbstregulierung der Medizin. Wenn sie die ärztliche Freiheit berührt, dann – wie Hill in Hinblick auf Archibald Cochrane kommentierte – als „internal challenge“ [45]. Hierzulande wird sie hingegen vielfach wahrgenommen als eine von der Gesundheitspolitik oktroyierte Maßnahme zur Einschränkung der ärztlichen Freiheit und zur bloßen Kostenreduktion.

2.6 Evidenzbasierte Medizin im Kontext aktueller Selbstverpflichtungen der ärztlichen Profession

Eine andere Bewegung, die Deutschland spät bzw. bis heute kaum erreicht hat, ist die des „New medical professionalism“ [46]. Die Spuren dieser Diskussion lassen sich in den USA, Kanada und in Großbritannien bis weit in die 1990er-Jahre zurückverfolgen. Ein wesentlicher Meilenstein ist die Charter on Medical Professionalism, die von verschiedenen nordamerikanischen und europäischen Ärztevereinigungen entwickelt und 2002 veröffentlicht wurde. Sie geht von einem informellen, aber vielfältig nachweisbaren Kontrakt zwischen Medizin und Gesellschaft aus. Professionalismus sei sein Fundament. In diesem Vertrag sichert „die Medizin“ „der Gesell-

schaft“ verlässliche Dienste in einem Bereich zu, der für jeden einzelnen Bürger wie für die Gesellschaft von fundamentaler Bedeutung ist, nämlich dem der Gesundheitserhaltung und Krankheitsbekämpfung. Andererseits gewährt „die Gesellschaft“ „der Medizin“, vor allem den Ärzten, besondere Privilegien, die sich im Wesentlichen den Begriffen Autonomie und Selbstbestimmung zuordnen lassen.

Der ärztliche Berufsstand kann (innerhalb gesetzlich normierter Grenzen) seine Aus-, Weiter- und Fortbildung sowie seine Standards und Qualitätssicherung selbst definieren bzw. organisieren, er verfügt über ein eigenes Recht mit eigener Gerichtsbarkeit und verwaltet sich selbst. Jeder Arzt (auch der angestellte) ist Angehöriger einer „Profession“, eines „freien Berufs“; er hat in den Grenzen der *lex artis* ein therapeutisches Privileg und genießt ein überdurchschnittliches Ansehen verbunden mit einem immer noch überdurchschnittlichen Einkommen.

Der Kontrakt ist informell, ungeschrieben. Dennoch kann er erodiert und gekündigt werden, vonseiten der Bürger, der Legislative und Exekutive, der Rechtsprechung, der Kostenträger – aber auch seitens der Profession und ihrer Mitglieder.

Das Lebenselixier des Kontrakts ist wechselseitiges Vertrauen. Die Charta stellt einen Versuch dar, dieses Vertrauen allgemein zu beschwören und durch verschiedene Selbstverpflichtungen der Profession zu befestigen. Insofern hat sie auch eine intraprofessionelle Funktion: eine durch diverse soziale Prozesse frustrierte Ärzteschaft soll auf ein einheitliches Verständnis ihres Berufsbildes und Berufsfeldes eingeschworen werden. Aus dieser Sicht erschließt sich der Sinn der Definition von „medical professionalism“ einer Arbeitsgruppe des Royal College of Physicians of London als „a set of values, behaviours, and relationships that underpin the trust the public has in doctors“ [47].

Die Charta formuliert drei fundamentale Prinzipien und 10 „commitments“ (Selbst-

verpflichtungen). Die drei Grundprinzipien sind: Vorrang des Wohlergehens der Patienten („primacy of patients' welfare“), Patientenautonomie und soziale Gerechtigkeit. Wenigstens sechs der Selbstverpflichtungen (in der Charta auch „professional responsibilities“ genannt) haben eine enge Verbindung zur EbM (vgl. Tab. 2.2).

Einen direkten Bezug zur EbM weisen auch der andernorts genannte Spirit of enquiry und die Forderung nach der Unterstützung von Forschungsvorhaben auf. Andere immer wiederkehrende Begriffe wie Mitgefühl („compassion“), Altruismus, Fürsorge („caring“), Respekt, allgemein: Achtung der Menschenrechte, Fürsprache/Eintreten für einzelne und Gruppen von Patienten („advocacy“), rechtzeitige Überweisung an andere Ärzte, Teamarbeit, besonnener Umgang mit Ressourcen und die in der Tabelle mit „-“ versehenen Verpflichtungen stehen für andersgeartete Leistungsansprüche an Ärzte und ihre Profession.

Ein Teil lässt sich dem in § 1 der ärztlichen Berufsordnung genannten Dienst an der Bevölkerung und damit sozialetischen Kategorien zuordnen. Ein anderer fokussiert auf Beziehungsaspekte des Behandlungsprozesses und die sie tragenden Werte, Pflichten und Tugenden.

Dem entspricht der Aufbau von § 70 SGB V, einer der Kernnormen dieses Sozialgesetzbuches. In Abs. 1 geht es einerseits um Fragen der Bedarfs- und Verteilungsgerechtigkeit (Satz 1), dann aber vor allem darum, dass die Versorgung der Versicherten „dem allgemein anerkannten Stand der medizinischen Erkenntnisse“ zu entsprechen habe, dass sie „ausreichend“ und „zweckmäßig“ sein solle, das „Maß des Notwendigen nicht überschreiten“ dürfe und in fachlich gebotener Qualität sowie wirtschaftlich erbracht“ werden solle (Sätze 1 und 2), also professionelle und ökonomische Vorgaben zu erfüllen habe. In Abs. 2 geht es dagegen um die Forderung einer „humane(n) Krankenbehandlung“.

Tab. 2.2: Ärztliche Selbstverpflichtungen und ihre Beziehung zur EbM

Selbstverpflichtung/Verantwortung	Beziehung zur EbM
Professional Competence	EbM als Methode lebenslanger Fortbildung
Honesty with Patients	Gemeinsame eb Entscheidungsfindung „evidence-informed patient choice“ *)
Patients' Confidentiality	–
Appropriate Relationships with Patients	–
Improving Quality of Care	Das (eb) Richtige tun, Fortschritt identifizieren und implementieren, Outcomes beobachten
Improving Access to Care	–
Just Distribution of finite Ressourcen	–
Scientific Knowledge	Evidenzproduktion fördern, Integrität der Forschung sichern, „Skeptikämie“
Managing Conflicts of Interest	Interessensgeleitete Verzerrungen (auch bei sich selbst) erkennen und kontrollieren
Professional Responsibilities	Definition von Standards, Zurückhaltung und Respekt gegenüber Kollegen, intraprofessionelle Kontrolle

eb = evidenzbasiert; *) Entwistle et al. 1998

Bisher hat sich die EbM mit besten Gründen vor allem auf das „Was“ der medizinischen Leistungserbringung konzentriert, weniger auf das „Wie“ und seine Wirtschaftlichkeit und noch weniger auf seine humanen Qualitäten. Das zeigt, dass EbM zuerst im Dienst einer klinischen Zweckrationalität steht. Es geht ihr zuerst um klinische Wirksamkeit, um klinischen Nutzen. Die die Zwecke legitimierenden Werte und Normen liegen grundsätzlich jenseits ihres Horizonts. So kann man in und mit der EbM sehr wohl über die zweckmäßigste Methode einer aktiven ärztlicher Sterbehilfe urteilen, nicht aber darüber, ob dieses Ziel ethisch gerechtfertigt werden kann. Ebenso kann man Evidenz dafür erarbeiten, ob eine patientenzentrierte(re) Behandlung deren klinische Ergebnisse oder die Patientenadhärenz oder -zufriedenheit verbessert. Aber das Ziel der Patientenzentrierung lässt sich darüber hinaus wiederum nicht aus der EbM ableiten; es verdankt sie anderen Quellen.

2.7 Ethik der Evidenzbasierten Medizin

Unter dieser Überschrift lassen sich ganz unterschiedliche Perspektiven eröffnen. Erstens kann das Kernkonzept der EbM ethisch reflektiert werden; dies scheint nicht sehr ergiebig, jedenfalls solange sich EbM als klinische Medizin versteht, die „nur“ die Auswahl ihrer Methoden und die Begründung ihrer Indikationen vorzugsweise durch Rückgriff auf die best erreichbare Evidenz aus systematischer Forschung reguliert und die darüber hinaus Ärzte verpflichtet (und es ihnen ermöglicht), mit dem allgemein anerkannten Stand der medizinischen Erkenntnisse Schritt zu halten. Dieses Programm scheint ethisch so gut wie unangreifbar [48]. Es verpflichtet die Medizin jedoch zu einer fortwährenden wissenschaftlichen Kontrolle ihrer Ergebnisse, also zur klinischen Forschung im engeren Sinn. Es verpflichtet damit auch die EbM, sich selbst auf ihre Folgen zu unter-

suchen. Führt eine evidenzbasierte klinische Praxis zu besseren Outcomes und zu einer wirtschaftlicheren Nutzung von Ressourcen (s. a. Kap. 17)?

Wenn evidenzbasierte klinische Medizin sich in ihren Handlungszielen und Handlungsschritten nicht wesentlich von herkömmlicher klinischer Medizin unterscheidet, dann entgeht sie auch keiner ihrer (bekannten) ethischen Spannungen wie z.B. dem Fall-Regel-Problem, der unaufhebbaren – wohl aber unnötig verschärfbaren – Asymmetrie in der Beziehung von Patient und Arzt oder der Frage nach der „Wahrheit am Krankenbett“. EbM ist nicht eo ipso patientenzentrierter als die klinische Medizin überhaupt.

Zweitens kann untersucht werden, welche ethisch relevanten Folgen sich aus dem Nicht-Gebrauch von EbM ergeben. Kann man sich eine Handlungswissenschaft Medizin ohne Rekurs auf eine methodisch-systematische Vergewärtigung ihrer Outcomes, also der Ergebnisse klinischer Forschung im engeren Sinne vorstellen? Wohl seit dem Ende des 19. Jahrhunderts nicht mehr! Selbst das Bundesverfassungsgericht fordert angesichts einer „lebensbedrohlichen oder regelmäßig tödlichen Erkrankung“ bei der „eine allgemein anerkannte, medizinischen Standards entsprechende Behandlung nicht zur Verfügung steht“, dass die gewählte Behandlungsmethode wenigstens „eine auf Indizien gestützte, nicht ganz fern liegende Aussicht auf Heilung oder wenigstens auf eine spürbare positive Einwirkung auf den Krankheitsverlauf versprechen“ müsse (1 BvR 347/98 vom 6.12.2005, Randnummer 64). Später ist dann ausdrücklich von Vergleichen zwischen behandelten und unbehandelten, aber in gleicher Weise erkrankter Personen, den daraus abgeleiteten Erfahrungen und von wissenschaftlichen Diskussionen die Rede.

Vor diesem Hintergrund werden vor allem diejenigen ethisch gefordert sein, die die EbM – aus welchen Gründen auch immer – ablehnen oder sich ihren Anforderungen zu

entziehen suchen (wie z.B. die „Cognition based medicine“ anthroposophisch orientierter Mediziner) [49; 50]. Sie finden sich unter Klinikern verschiedenster Disziplinen und Psychotherapeuten ebenso wie unter Ärztefunktionären, Sozialrichtern und Anhängern besonderer Therapierichtungen. Ihnen schwebt eine auf unmittelbaren persönlichen Erfahrungen und jeweils besonderen theoretischen Einsichten gegründete Individualmedizin vor. Verdrängt wird die praktische Unmöglichkeit einer solchen Medizin; jeder Arzt folgt, beobachtet man ihn einige Wochen, bestimmten Handlungsschemata; keiner erfindet sein Vorgehen in jeder klinischen Situation neu. Selten lassen sich die Handlungsabfolgen allerdings so gut begründen wie die in der EbM – etwa in Leitlinien – empfohlenen.

Drittens sind ethische Probleme zu reflektieren, die sich mithilfe der EbM vermeiden lassen. Die EbM ist von Klinikern zur Minderung auch ethischer Spannungslagen entwickelt worden (u.a. verspätete oder verfrühte Einführung neuer Methoden, Identifizierung dringenden Forschungsbedarfs oder redundanter Forschung). EbM ist ein Werkzeug, um verschiedene (Selbst-)Verpflichtungen des Arztes und seiner Profession einzelnen Kranken, Krankengruppen und der Bevölkerung bzw. Gesellschaft gegenüber besser erfüllen zu können (s. Tab. 2.2).

Bevor man sich – viertens – näher mit den durch die EbM geschaffenen oder verschärften Spannungslagen beschäftigt, ist zu klären, ob man die Theorie der EbM oder ihre Praxis meint. Der Vorwurf etwa der „Kochbuchmedizin“ träfe dann nicht, wenn man in der Klinik überwiegend auf EbM-PropONENTEN stieße, die Leitlinienempfehlungen flexibel anwenden, unter Berücksichtigung der klinischen Lage und der Präferenzen der Patienten (s. Abb. 2.1). EbM würde zur Karikatur ihrer selbst, wenn Ärzte Leitlinienempfehlungen in der klinischen Situation als Richtlinien auffassen und blind exekutieren

würden; EbM würde die Patientenbehandlung lähmen, wenn man sich klinisch allein an der angeblich besten vorstellbaren Evidenz aus RCTs, nicht aber an der problemadäquaten und erreichbaren Evidenz orientieren wollte. Manche Kritik an der EbM scheint mit ihrer Wirklichkeit wenig zu tun zu haben, konzentriert sich auf ihre vermeintlich unreflektierte Theorie [22] und lebt vom Abwehrmechanismus der Projektion.

Dennoch hat EbM verschiedene spezifische Risiko- und Missbrauchspotenziale; einige seien genannt:

- ▲ Regelmäßig orientieren entscheidungsrelevante Therapiestudien auf mittlere Effekte. Die Varianz um die herausgestellte zentrale Tendenz (Überlebensrate, Relatives Risiko, Mittelwertdifferenz) und ihre Determinanten wird übersehen. Ist die Prognose im Mittel ungünstig, könnten Hoffnungen und Anstrengungsbereitschaft beeinträchtigt werden, aufseiten des Arztes wie des Patienten. Patienten glauben (wie auch Ärzte), das Unwahrscheinliche erreichen zu können. Es wäre inhuman und oft schädlich, ihnen diesen Glauben unter Hinweis auf mittlere Prognosen zu nehmen.
- ▲ Eine Konsequenz einer Fixierung auf RCTs könnte es auch sein, nur noch solche Patienten aktiv zu behandeln, die deren enge Einschlusskriterien vollständig erfüllen und damit anderen Patienten eine womöglich aussichtsreiche Therapie vorzuhalten. Inwieweit können Studienergebnisse auf weitere Patientengruppen generalisiert werden, inwieweit dürfen individuelle Patienten evidenzbasierten Studienergebnissen unter- bzw. zugeordnet werden? Dies ist eine Frage, die besonders auch die Gruppen interessiert, die die Qualität von Evidenz und Empfehlungsstärken graduieren möchten [13; 51].
- ▲ Ungeklärt ist, in welchen klinischen Situationen welche Evidenz- und Empfehlungsstärken zu fordern sind. Im Bereich der Primär- und Sekundärprävention (und auch der Rehabilitation und Pflege) sind die Anforderungen aktuell deutlich lockerer als im Bereich der Therapie. Das zitierte Urteil des Bundesverfassungsgerichts hat darauf aufmerksam gemacht, dass die Hürden für die Wahl und Erstattung auch alternativer Methoden im Fall lebensbedrohlicher Krankheiten sehr niedrig sind. Höher liegen sie nach der Rechtsprechung des Bundessozialgerichts dort (B 1 KR 37/00 R, 19.3.2002), wo ein so genannter Off label use zugelassener Medikamente bei schwerwiegender Krankheit zu rechtfertigen ist, wenn keine andere Therapie verfügbar ist und „aufgrund der Datenlage die begründete Aussicht besteht, dass mit dem betreffenden Präparat ein Behandlungserfolg erzielt werden kann“. Und noch einmal anspruchsvoller ist der Gemeinsame Bundesausschuss in der Beurteilung „gewöhnlicher“ neuer und alter Methoden.
- ▲ Ein aktives Übersehen der Grenzen und Defizite von RCTs geht oft mit dem Gefühl moralischer Überlegenheit einher und verführt dazu, EbM als Waffe einzusetzen. Man kann dann z.B. das Fehlen von RCT-Evidenz grundsätzlich mit einer Evidenz für das Fehlen von Wirksamkeit und Nutzen verwechseln; man kann, auch aus ökonomischen Gründen, mit dem Hinweis auf fehlende RCTs neue und alte Methoden abwehren. Andererseits kann man mit RCTs in Evidenzfällen hineingeraten: diese werden gern in Form hochrangig (und selektiv) publizierter Ergebnisse von in verschiedener Hinsicht begrenzten RCTs aufgestellt. Oft handelt es sich um die allererste oder zweite Veröffentlichungen zu einer Methode (Optimism-Bias), [52]. Die Entzauberung des Verfahrens setzt regelmäßig mit Re-Analysen der ersten Publikationen, mit späteren Veröffentlichungen und Daten aus Anwendungsbeobachtungen ein.

- ▲ Ein Missbrauch von EbM läge nahe, wenn man glaubte, aus vorliegender Evidenz unmittelbar rechtlich bindende Entscheidungen ableiten zu müssen. Es ist auf dem Unterschied von Evidenz, evidenzbasierter Empfehlung und faktischer Entscheidung zu beharren. So wird etwa die Screening-Mammographie in Deutschland eingeführt (ähnlich die Disease-Management-Programme), ohne dass sich die dahinter stehenden politischen Institutionen und (Selbst-)Verwaltungen auf überzeugende empirische Evidenz hätten berufen können. In der internationalen Literatur war im Zusammenhang mit dem Mammographie-Screening von „agonies of evidence“ (BMJ 10.11.2001) und einer „crisis for evidence-based medicine?“ (Annals Internal Medicine 3.9.2002) zu lesen. Letztere ist bis heute nicht beigelegt [53]. Dennoch ist es legitim und auch von der EbM hinzunehmen, wenn demokratisch hierzu ermächtigte Institutionen abweichende Entscheidungen treffen; sie können sich dann nur nicht auf Evidenz berufen, sondern müssen ihre anderen Gründe offen legen. Evidenz trifft keine Entscheidungen, sie unterstützt eine empirisch-zweckrational begründbare Entscheidungsfindung. Dies gilt besonders für Entscheidungen auf der Systemebene, wo regelmäßig neben der Bedarfsgerechtigkeit auch die Verteilungsgerechtigkeit (und viele andere politische und ökonomische Faktoren) zu berücksichtigen sind.
- ▲ Für die zu treffenden Entscheidungen sind auch die in der EbM weitgehend ausgeblendeten Wertentscheidungen von Bedeutung. So sind Situationen denkbar, in denen – aus Gründen knapper Ressourcen – nicht alle evidenzbasierten Indikationen verwirklicht werden können. Welche sollen dann vorgehen, welche zurückstehen? Dabei werden die Häufigkeit, Schwere und Gefährlichkeit der Erkrankungen ebenso zu berücksichtigen sein wie die Dignität

der Ziele, die Nutzen- und Schadenspotenziale und die Kosten-Nutzen-Relation der auf sie bezogenen Leistungen, deren Notwendigkeit bzw. Ersetzbarkeit und soziale Akzeptanz. Auch ethische und juristische Normen fallen ins Gewicht. Für viele Detailfragen wird man Evidenz beibringen können und müssen – ohne dass damit schon Abwägungen angestellt und Entscheidungen getroffen wären.

Schließlich gibt es (fünftens) eine Ethik der der EbM zugrunde liegenden (klinischen) Forschung und der Publikation, Analyse wie Synthese ihrer Ergebnisse. Allgemein ist eine Abnahme der Beteiligungsraten an verschiedenen Arten epidemiologischer und klinischer Forschungsvorhaben zu verzeichnen. Um so wichtiger ist es, das Vertrauen von Patienten und Probanden in die klinische Forschung, ihre Fragestellungen, Studiendesigns, Studiendurchführung, Auswertungen und Veröffentlichungen zu erhalten. Klinische Forschung könnte vollständig zum Erliegen kommen, wenn die Kranken (und die gesunden Probanden) ihre freiwillige Mitarbeit aufkündigten. Dies macht es so wichtig, alle Forschungsvorhaben, auch die von Nicht-Ärzten (Naturwissenschaftler, Psychologen, Krankengymnasten usw.) geplanten, einer Ethikkommission zur Beratung und objektivierenden Chancen-Risiken-Abwägung vorzulegen. Die Richtlinie der Europäischen Gemeinschaft 2001/20/EG sieht eine Ethikkommission als ein unabhängiges Gremium, „dessen Aufgabe es ist, den Schutz der Rechte, die Sicherheit und das Wohlergehen von an einer klinischen Prüfung teilnehmenden Personen zu sichern und diesbezüglich Vertrauen der Öffentlichkeit zu schaffen“. Die Prüfung konzentriert sich auf die wissenschaftliche Qualität der Studie, auf die Einhaltung rechtlicher Vorgaben und auf ethische Fragen im engsten Sinn (u.a. Patienteninformation und -einwilligung, Einwilligungsfähigkeit, Risiko-Nutzen-Abwägung).

2.8 Schlussbemerkung

In Deutschland ist die EbM, wie angedeutet, von der Mehrzahl der klinisch tätigen Ärzte als „Revolution von oben“ erlebt worden. Eine Integration ihres Konzepts, ihrer Methoden und Techniken in den klinischen Alltag scheint nur an wenigen Stellen gelungen.

Die Vermittlung ihrer theoretischen und empirisch-wissenschaftlichen Grundlagen einschließlich ihrer Geschichte und Ethik sollte die Akzeptanz der EbM im klinischen Alltag unterstützen können. EbM ist der ärztlichen Profession weniger fremd, als es ihre Kritiker unterstellen. Würde sie stärker als „selbst“ erkannt, könnten sich manche Abstoßungsreaktionen vermutlich kontrollieren, dämpfen oder auch vermeiden lassen.

Literatur

- [1] Muir Gray JA (1997) Evidence-based Healthcare. How to make health policy and management decisions. Churchill Livingstone, Edinburgh
- [2] Callahan D (2002) Ziele der Medizin: Welche neuen Prioritäten müssen gesetzt werden? In: Allert G, Ziele der Medizin, 22–81. Schattauer, Stuttgart
- [3] Hartmann F (1987) Leben in bedingtem Gesundheit. In: Schaefer H, Schipperges H, Wagner G, Präventive Medizin, 233–253. Springer, Berlin
- [4] Eulenburg A, Stolberg M (Ed) (1894) Real-Encyclopädie der gesammten Heilkunde. Urban und Schwarzenberg, Wien
- [5] Sackett DL et al., Evidence based medicine: what it is and what it isn't. BMJ. (1996), 312, 71–72
- [6] Haynes RB, Devereaux PJ, Guyatt GH, Physicians' and patients' choices in evidence based practice. BMJ. (2002), 324, 1350
- [7] Spitzer WO, Clinical epidemiology. J Chronic Dis. (1986), 39, 411–415
- [8] Schwartz D, Lellouch J, Explanatory and pragmatic attitudes in therapeutical trials. J Chronic Dis. (1967), 20, 637–648
- [9] Willich SN, Randomisierte kontrollierte Studien: Pragmatische Ansätze erforderlich. Dtsch Arztebl. (2006), 103, A–2524–30
- [10] Tunis SR, Stryer DB, Clancy CM, Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. JAMA. (2003), 290, 1624–1632
- [11] Gartlehner G et al., A simple and valid tool distinguished efficacy from effectiveness studies. J Clin Epidemiol. (2006), 59, 1040–1048
- [12] Treweek S et al., The Trial Protocol Tool: The PRACTIHC software tool that supported the writing of protocols for pragmatic randomized controlled trials. J Clin Epidemiol. (2006), 59, 1127–1133
- [13] Atkins D et al., Grading quality of evidence and strength of recommendations. BMJ. (2004), 328, 1490–1497
- [14] Guyatt G et al., Grading strength of recommendations and quality of evidence in clinical guidelines: report from an american college of chest physicians task force. Chest. (2006), 129, 174–181
- [15] Deutsche Forschungsgemeinschaft (1999) Klinische Forschung: Denkschrift. Wiley-VCH, Weinheim
- [16] Shulman LE, Clinical research 1996: stirrings from the academic health centers. Acad Med. (1996), 71, 362–3, 398
- [17] Martini P (1951) Experimentelle und therapeutisch-klinische Forschung, Kritik und Ethik. Festvortrag zur Eröffnung der Deutschen Therapiewoche 1950. Med.Wiss.Verl., Karlsruhe
- [18] Bonell C et al., Assessment of generalisability in trials of health interventions: suggested framework and systematic review. BMJ. (2006), 333, 346–349
- [19] Bortz J, Döring N (2002) Forschungsmethoden und Evaluation. Springer, Berlin
- [20] Hartmann F (1989) Medizin – eine Wissenschaft aus eigenem Recht? In: Rössler D, Waller HD, Medizin zwischen Geisteswissenschaft und Naturwissenschaft, 21–44. Attempto, Tübingen
- [21] Naunyn B, Aerzte und Laien. Dtsch Revue. (1905), 30, 343–355
- [22] Raspe H (2005) Konzept und Methoden der Evidenz-basierten Medizin: Besonderheiten, Stärken, Grenzen, Schwächen und Kritik. In: Kolmar M, Freudenberg U, Marckmann G, Effizienzdenken und moralische Verpflichtung im Gesundheitswesen. Studienbrief, 198–246. FernUniversität, Hagen

- [23] Hartmann F (1994) Gedanken zum therapeutischen Imperativ. In: Meier J, Menschenbilder-Philosophie im Krankenhaus, Olms, Hildesheim
- [24] Katsch G (1958) Der therapeutische Imperativ des Arztes. Lehmann, München
- [25] Sackett DL, Clinical epidemiology. what, who, and whither. *J Clin Epidemiol.* (2002), 55, 1161–1166
- [26] Sackett DL, Clinical epidemiology. *Am J Epidemiol.* (1969), 89, 125–128
- [27] Haynes RB et al. (2005) Clinical epidemiology. How to do clinical practice research. Lippincott Williams and Wilkins, London
- [28] Stoll S, Roelcke V, Raspe H, Gibt es eine deutsche Vorgeschichte der Evidenz-basierten Medizin? Methodische Standards therapeutischer Forschung im beginnenden 20. Jahrhundert und Paul Martinis „Methodenlehre der therapeutischen Untersuchung“ (1932). *Dtsch Med Wochenschr.* (2005), 130, 1781–1784
- [29] Martini P (1932) Methodenlehre der therapeutisch-klinischen Forschung. Springer, Berlin
- [30] Tröhler U (2000) To improve the evidence of medicine. The 18th century British origins of a critical approach. Royal College of Physicians of Edingburgh
- [31] Schulz S et al. (2006) Geschichte, Theorie und Ethik der Medizin. Suhrkamp, Frankfurt/Main
- [32] Matthews JR (1995) Quantification and the quest for medical certainty. Princeton Univ. Pr., Princeton
- [33] Marshall G et al., Streptomycin treatment of pulmonary tuberculosis. *BMJ.* (1948), 769–782
- [34] Bingel A, Über Behandlung der Diphtherie mit gewöhnlichem Pferdeserum. *Dtsch Archiv Klin Med.* (1918), 125, 284–332
- [35] Pullar T, Capell HA, Sulphasalazine: a „new“ antirheumatic drug. *Br J Rheumatol.* (1984), 23, 26–34
- [36] Stoll S, Klinische Forschung und Ethik bei Paul Martini. *Z Arztl Fortbild Qualitatssich.* (2003), 97, 675–679
- [37] Hill AB (1937) Principles of medical statistics. *Lancet*, London
- [38] Feinstein AR (1985) Clinical epidemiology. The architecture of clinical research. Saunders, Philadelphia
- [39] Weiss NS (1986) Clinical epidemiology: The study of outcome of illness. Oxford Univ. Pr., Oxford
- [40] Kramer MS (1988) Clinical epidemiology and biostatistics. Springer, Berlin
- [41] Sackett DL, Haynes RB, Tugwell P (1985) Clinical epidemiology. Little, Brown and Co., Boston
- [42] Ollenschläger G, Kunz R, Jonitz G, Evidence-Based Medicine: Wege zu einer rationaleren Medizin. *Dtsch Arztebl.* (1998), 95, A–267–70
- [43] Donner-Banzhoff N, Allgemeinmedizin an der Universität Toronto: Hausärzte forschen für die Praxis. *Dtsch Arztebl.* (1996), 93, A–2616–7
- [44] Raspe H, Evidence based medicine: Modischer Unsinn, alter Wein in neuen Schläuchen oder aktuelle Notwendigkeit? *Z Arztl Fortbild (Jena).* (1996), 90, 553–562
- [45] Hill GB, Archie Cochrane and his legacy. An internal challenge to physicians' autonomy? *J Clin Epidemiol.* (2000), 53, 1189–1192
- [46] Irvine D, Doctors in the UK: their new professionalism and its regulatory framework. *Lancet.* (2001), 358, 1807–1810
- [47] Royal College of Physicians (2005) Doctors in society. Medical professionalism in a changing world. RCP, London
- [48] Raspe H, Ethische Implikationen der Evidenz-basierten Medizin. *Dtsch Med Wochenschr.* (2002), 127, 1769–1773
- [49] Kiene H (2001) Komplementäre Methodenlehre der klinischen Forschung. Springer, Berlin
- [50] Raspe H, Cognition-based Medizin aus der Sicht der Evidenz-basierten Medizin. *Z Arztl Fortbild Qualitatssich.* (2005), 99, 295–300
- [51] Atkins D et al., Systems for grading the quality of evidence and the strength of recommendations II: pilot study of a new system. *BMC Health Serv Res.* (2005), 5, 25
- [52] Chalmers I, Matthews R, What are the implications of optimism bias in clinical research? *Lancet.* (2006), 367, 449–450
- [53] Goetzsche PC, Nielsen M, The Cochrane Collaboration (2006) Screening for breast cancer with mammography (Review). Wiley, New York

3 Evidenzbasierte Medizin macht Spaß oder: Das Richtige richtig machen

Wolfgang A. Blank und Günther Egidi

3.1 Vorsicht: EbM kann süchtig machen!

Erfolgserebnisse sind für viele niedergelassene Ärzte oft spärlich gesät. Der zeitlich straff organisierte Arbeitstag, die Fülle verschiedenster Beratungsanlässe und die beim Arzt-Patienten-Kontakt notwendige Empathie lassen den Praxisalltag oft zum Stress ausarten und geben wenig Raum für die (selbst-)kritische Überprüfung des eigenen Handelns [1].

Situationen, in denen es dem Arzt unter diesen Bedingungen gelingt, ein kniffliges Problem zu lösen, einen gefährlichen Verlauf abzuwenden oder einfach nur das sichere Gefühl zu haben, das medizinisch Richtige für seinen Patienten getan zu haben, sind in der Regel rar. Dabei könnten diese Situationen einen Großteil ärztlicher Zufriedenheit ausmachen [2].

3.1.1 Es fehlt nicht an Gelegenheiten!

Gerade zu den Problemen der regelmäßig in der Praxis behandelten chronisch Kranken mit arterieller Hypertonie, Diabetes mellitus oder Herzinsuffizienz liegen umfangreiche Studien und Leitlinien mit guter Evidenz vor [3; 4]. Dieses wissenschaftlich fundierte Wissen praktikabel in den täglichen Alltag zu integrieren, kann die Arbeitszufriedenheit beträchtlich erhöhen [5].

3.2 EbM in der Hausarzt-Praxis – Szenario

Frau Stern arbeitet als Hausärztin in einem ländlichen Gebiet. Vor 10 Tagen hat sie eine 72-jährige Patientin mit einer akut aufgetretenen Hemiparese in die neurologische Klinik der nahe gelegenen Kreisstadt eingewiesen. Schon vor einem Jahr hatte die Patientin eine TIA erlitten und nahm seitdem regelmäßig 100 mg ASS ein. Auch dieses Mal ist es glücklicherweise bei einer TIA geblieben – die neurologischen Defizite bildeten sich komplett zurück.

Die Klinik empfahl im Entlassungsbrief die Rezidivprophylaxe mit 75 mg Clopidogrel, da ASS nicht wirksam gewesen sei.

3.3 Im Trott gefangen?

3.3.1 Situation des Niedergelassenen

Der Arzt ist bestrebt, seine Patienten nach bestem Wissen und Gewissen zu betreuen. Unter den Prämissen der Gesundheitspolitik (§12 SGB V: wirtschaftlich, ausreichend, notwendig und zweckmäßig) greift er sowohl auf seine Erfahrungen als auch auf aktuell zur Verfügung stehende Informationen zurück.

Szenario

Frau Stern ist eine sorgfältige Ärztin und möchte ihrer Patientin einen dritten Schlaganfall ersparen. Doch sie zweifelt, ob mit Clopidogrel der drohende Apoplex eher verhindert werden kann als mit der weiteren ASS-Gabe. Auch die Tages-Therapiekosten von 2,43 € für 75 mg Clopido-

grel gegenüber 0,03 € bei 100 mg ASS motivieren die von Arzneimittelregressen bedrohte Kassenärztin, ihre Therapieentscheidung gründlich zu bedenken. Frau Stern verschreibt der Klinikempfehlung folgend eine kleine Medikamentenpackung, um Zeit zu gewinnen. In ihrer Landarztpraxis mit 1500 Scheinen findet sie kaum Zeit, sofort eine ausführliche Literaturrecherche durchzuführen.

3.3.2 Das Wissen des Arztes

Als Grundlagen dienen ihm primär die Erfahrungen und Kenntnisse aus dem Studium und der Zeit im Krankenhaus. Im weiteren beruflichen Leben kommen die Lektüre einschlägiger Zeitschriften und Fortbildungen hinzu. Diese enthalten oft nicht belegte oder tendenziös aufbereitete Informationen [6]. Darüber hinaus greift er auf seine im Laufe seines weiteren Berufslebens gesammelten persönlichen Erfahrungen zurück. Hieraus resultiert das Wissen, das der ärztlichen Beratung und Therapie zugrunde liegt.

3.3.3 Die Sicherheit der ärztlichen Therapie

Die strikte Therapie nach Leitlinien garantiert nicht, dass das Vorgehen des Arztes dem Patienten nützt [7]. Auch Empfehlungen, den Lebensstil zu ändern, sind häufig – obwohl sie einleuchtend klingen – nicht belegt. Hier baut sich ein Spannungsfeld zwischen den Erfahrungen des Arztes, übernommem Wissen, dem Anspruch, das Beste für seinen Patienten erreichen zu wollen, und der harten wissenschaftlichen Realität auf. Eine Zwickmühle, aus der sich viele Ärzte nicht aus eigener Kraft befreien können.

3.4 Hilfsmittel EbM

Evidenzbasierte Medizin ist eines der Hilfsmittel, um dieses Spannungsfeld aufzulösen. Valide Studien bilden Tatsachen ab, auf die sich erfolgreiche Therapieentscheidungen sinnvoll gründen lassen. EbM wird so zum Handwerkszeug, das – richtig und geschickt angewandt – die Behandlung für Arzt und Patienten zu einem Erfolgserlebnis werden lassen kann.

3.5 Viele Vorteile

3.5.1 Vorteil für den Arzt

Der Arzt gewinnt Sicherheit in seinen Entscheidungen. Eine Behandlung wird auf der Grundlage unabhängiger und gesicherter – nicht nur vermuteter – Vorteile durchgeführt. So können bewusst oder unbewusst in die ärztliche Therapieentscheidung eingebundene Fremdeinflüsse ausgeschlossen werden. Die Gefahr, zum Handlanger eigennütziger Interessen anderer, am Gesundheitssystem beteiligter Gruppen zu werden [8], kann so vermieden werden. Eine Erleichterung für den ethisch korrekten Arzt, denn die medizinisch-wissenschaftlich richtige Entscheidung (unter Beachtung des Patientenwillens und der Erfahrung des Arztes) steht wieder im Mittelpunkt des ärztlichen Handelns. Ein weiterer Nebeneffekt: Auf diese Weise werden nicht selten auch die Ressourcen des öffentlichen Gesundheitswesens geschont.

3.5.2 Vorteil für den Patienten

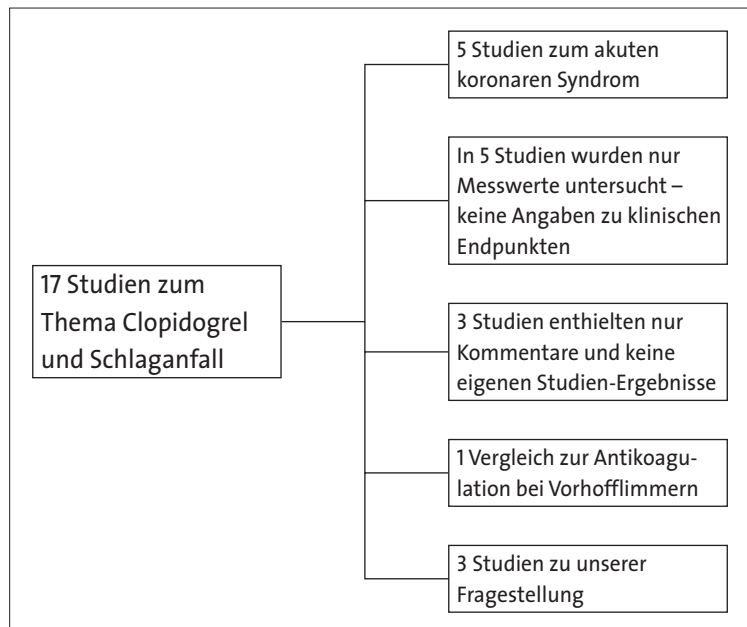
Der Patient wird – beispielsweise durch die Boulevardpresse – mit medizinischen Informationen überschüttet [9]. Die verwirrende Vielfalt angeblich Heil bringender Behandlungen – auch im Internet – verunsichert

nicht selten Patienten in ihrer Entscheidung. Eine qualifiziert durchgeführte Studie mit für den Patienten wichtigen und nachvollziehbaren Ergebnissen kann ihm – bei patientenorientierter Darstellung durch den Arzt – Sicherheit in seiner persönlichen Entscheidung geben. Auch die Wertigkeit verschiedener Maßnahmen (Umstellung der Lebensgewohnheiten oder Medikamente) kann klarer dargestellt werden. Gemeinsam getroffene Entscheidungen erreichen dadurch eine höhere Adhärenz beim Patienten [10].

3.5.3 Vorteil für Studenten und Berufsanfänger

Der undurchdringlich erscheinende Dschungel der verschiedenen Therapieempfehlungen und die unterschiedliche Handhabung derselben in den jeweiligen Klinikabteilungen kann durch die Verwendung evidenzbasierter Literatur gelichtet werden. Korrekt ermittelte und relevante Fakten sind Argumente, die auch in eingespielten Klinikteams nicht einfach wegdiskutiert werden können.

Abb. 3.1: Ergebnisse Internetrecherche Clopidogrel (Pubmed-Suche durchgeführt am 4.7.2006)



So werden selbst Berufsanfänger in der Therapie verschiedener Erkrankungen schnell den Überblick gewinnen [11].

3.5.4 Vorteil für Qualitätszirkel

Der kollegiale Austausch im Ärztekreis gewinnt an Tiefe, wenn valide Fakten der Diskussion zugrunde liegen. Das Interesse an aktuellen Recherchen auf der Basis wissenschaftlich fundierter Ergebnisse entsteht fast von selbst. So können Vergleichsstudien herangezogen und im Kollegenkreis die Praxis-tauglichkeit der EbM-Empfehlungen überprüft werden.

Der Qualitätszirkel als Instrument des qualifizierten kollegialen Austausches hat so die Chance, die Einfügung externer Evidenz zusätzlich zu den Faktoren Arzterfahrung und Patientenwunsch zu gewährleisten.

Szenario

Die Lösung bietet sich Frau Stern in Form ihres Qualitätszirkels an. Seit einiger Zeit nutzen hier benachbarte Hausärzte die Technik, jeweils ein

bis zwei in der täglichen Praxis offen bleibende Fragen zu notieren, die dann von den Kollegen in der Runde gemeinsam bearbeitet werden. Findet sich keine Lösung, übernimmt einer der Kollegen die Aufgabe, bis zum nächsten Treffen eine Antwort zu finden.

Alle Mitglieder des Qualitätszirkels beherrschen die PubMed-Recherche. Sie haben sich erst vor kurzem in Internet-Recherchen schulen lassen.

Ihr Kollege Müller interessiert sich für die Sekundärprophylaxe nach Apoplex und verspricht, bis zum nächsten Qualitätszirkel Frau Sterns Frage nachzugehen.

Unter den Suchbegriffen „Clopidogrel“ und „cerebrovascular accident“ mit der Begrenzung auf randomisierte kontrollierte Untersuchungen an Menschen landet er unter <http://www.ncbi.nlm.nih.gov/entrez/> siebzehn Treffer (s. Abb. 3.1).

3.6 Das Werkzeug der EbM

3.6.1 Der tatsächliche Nutzen

Primäre Entscheidungsgrundlage für die Wirksamkeit einer diagnostischen oder therapeutischen Maßnahme muss sein, inwieweit ein potenzielles Risiko (hier der Apoplex) durch eine Intervention reduziert wird.

Szenario

Das Ergebnis von zwei der drei Studien zu unserer Frage zu Clopidogrel ist, dass die gleichzeitige Gabe von Clopidogrel und ASS bei rezidivierenden zerebralen Ereignissen eher schadet als nützt [12; 13]. Das ist zwar sehr interessant, beantwortet aber nicht die Frage, ob bei Rezidiv-Insult ASS durch Clopidogrel ersetzt werden muss.

Schnell wird klar: Nur die dritte, älteste Untersuchung, die CAPRIE-Studie [14] kann die Frage von Herrn Müller beantworten. In dieser großen Studie wurde der Nutzen von Clopidogrel im Vergleich zu ASS zur Verhinderung von Schlaganfall-Rezidiven untersucht. 19 185 Pa-

tienten erhielten eine der beiden Substanzen. 509 von 9599 mit Clopidogrel und 546 von 9586 mit ASS behandelten Patienten erlitten während der 1,91 Jahre dauernden Beobachtung einen Insult.

Einen Haken dieser Studie findet Herr Müller allerdings ziemlich schnell: Der Schlaganfall war kein vor Studienbeginn klar definierter Untersuchungs-Gegenstand (primärer Endpunkt). Der beobachtete gemischte primäre Endpunkt setzte sich zusammen aus Schlaganfall, Herzinfarkt und kardiovaskulärem Tod. Statistisch gesehen ist es unzulässig, die Studie einzig mit Blick auf die Verminderung eines Schlaganfalles zu sehen, weil sie darauf nicht angelegt war. Herr Müller darf das Ergebnis nur bezüglich der gesammelten Endpunkte Schlaganfall, Herzinfarkt und kardiovaskulärer Tod interpretieren. Da es sich insgesamt um vaskuläre Ereignisse handelt und die Tendenz bei allen einzelnen Endpunkten ähnlich ist, verfolgt Herr Müller diese Studie jedoch weiter.

In der CAPRIE-Studie traten unter ASS bei 1021 von 9586 (entsprechend 5,83% Ereignisse jährlich), unter Clopidogrel bei 939 von 9599 Fällen (entsprechend 5,32% Ereignisse jährlich) diese ischämischen Ereignisse auf.

Das bedeutet konkret: Die Wahrscheinlichkeit, mit Clopidogrel kein vaskuläres Ereignis zu erleiden, beträgt 8,75% (vgl. 3.7) – eine Zahl, die Kollege Müller dem Qualitätszirkel beim nächsten Treffen, von der Überlegenheit des Clopidogrel gegenüber dem ASS überzeugt, präsentiert. Der Einsatz der teureren Substanz scheint sich zu rentieren.

3.7 Zahlenspiele

Den Arzt und seine Patienten interessiert, in welchem Ausmaß das Risiko eines unerwünschten Ereignisses durch eine gezielte Maßnahme (hier: Gabe von Clopidogrel) reduziert wird.

Leider ist diese Einschätzung der Risikoreduktion für den Betrachter nicht leicht.

Denn es macht einen großen Unterschied, ob der Anteil derer, die unter Clopidogrel weniger ischämische Ereignisse erlitten, absolut oder relativ betrachtet wird. Im ersten Fall geht es um das Konzept der absoluten Risiko-Differenz, im zweiten um das des relativen Risikos bzw. der Relativen Risiko-Reduktion.

Es werden zwei absolute Risiken zueinander in Beziehung gesetzt – nur auf völlig unterschiedliche Weise. Die gleichen Zahlen, und doch – je nach Interpretation – zwei deutlich unterschiedliche Reduktionen der Risiken? Die Lösung ist einfach, aber hochbrisant: Es kommt darauf an, ob man die Differenz zwischen den Risiken (Absolute Risiko-Reduktion) oder das Verhältnis der Risiken zueinander (Relative Risiko-Reduktion) betrachtet.

Als **absolutes Risiko** wird die Wahrscheinlichkeit eines als ungünstig beurteilten Ereignisses ausgedrückt, das in einer definierten Gruppe und in einem definierten festgelegten Zeitraum auftritt. Das ungünstige Ereignis (Insult, Herzinfarkt oder vaskulärer Tod) trat in der Clopidogrel-Gruppe in 5,32 %, in der ASS-Gruppe in 5,83% der Fälle auf. Die Differenz zwischen den Risiken 5,83% (ASS) und 5,32% (Clopidogrel) beträgt 0,51%. Fazit: In der Clopidogrel Gruppe haben 0,51% weniger Patienten einen Insult erlitten. **Dies entspricht der Absoluten Risiko-Reduktion (ARR).**

Als **relatives Risiko** wird dagegen das Verhältnis der in den verschiedenen Gruppen aufgetretenen Risiken zueinander betrachtet ($0,0532/0,0583 = 0,9125$).

Die **Relative Risiko-Reduktion** durch die Gabe von Clopidogrel, die den Betrachter interessiert, ergibt sich somit aus dem Verhältnis der absolut weniger aufgetretenen ungünstigen Ereignisse in der Gruppe mit einer Maßnahme (Clopidogrel: 0,51% weniger Ereignisse) zu dem absoluten Risiko in der Gruppe ohne diese Maßnahme (ASS: 5,83% kardiovaskuläre Ereignisse). **Diese Relative**

Risiko-Reduktion (RRR) errechnet sich über die einfache Prozentrechnung: $0,0051/0,0583 = 0,0875$.

Je nach Betrachtungsweise und individuellen Interessen kann man also die Verminderung des Risikos mit mageren absoluten 0,51% oder eher beeindruckenderen relativen 8,75% angeben – bei gleich bleibender Zahlengrundlage.

Bei kleinen absoluten Risiken ist die Relative Risiko-Reduktion fast immer eindrucksvoller als die Absolute Risiko-Reduktion. Wenn nun die Wirksamkeit einer Intervention herausgestellt werden soll, wird fast zwangsläufig auf den höheren Wert der Relativen Risiko-Reduktion zurückgegriffen werden. Aus diesem Grund spielt letztere in der Pharmawerbung die hervorragende Rolle. Das ist publizistisch zwar nachvollziehbar, aber wissenschaftlich höchst zweifelhaft.

Praktische Bedeutung erhält die Unterscheidung zwischen relativem und absolutem Risiko vor allem im hausärztlichen, so genannten Niedrig-Prävalenz-Bereich. Unter der Vielzahl von Patienten finden sich kritische Ereignisse eher selten. So sind auch absolutes Risiko und Absolute Risiko-Reduktion in der Regel ausgesprochen niedrig. Für die Überlegung des Arztes, eine individuelle Behandlung einzuleiten, ist in der Praxis aber gerade dieses – meist niedrig ausfallende – absolute Risiko von Bedeutung.

Szenario

Ganz überzeugt ist der Qualitätszirkel deshalb von Herrn Müllers Schlussfolgerung nicht. Frau Stern wirft ein, dass es sich angesichts der großen Anzahl beobachteter Patienten nur um einen eher geringen Nutzen handele. Sie rechnet mit den absoluten Zahlen: Jährlich kam es nur bei ca. jedem 20. Patienten (5,83% bzw. 5,32%) zu einem ischämischen Ereignis.

Bezieht man den Effekt des Clopidogrel von relativen 8,75% (Relative Risiko-Reduktion/RRR) auf alle untersuchten Patienten, schmilzt der Vorteil von relativen 8,75% auf absolute

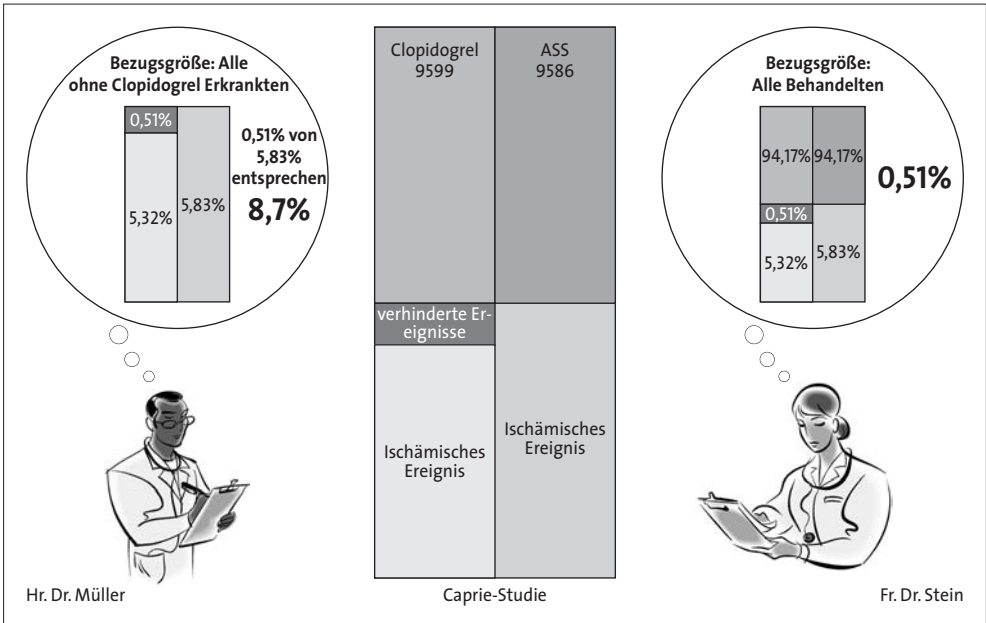


Abb. 3.2: Studienergebnis mit beiden Ärzten und Interpretation nach RRR und ARR

0,51% (Absolute Risiko-Reduktion = ARR). Die Gruppe überlegt nun, wie sich diese Zahlen in der täglichen Beratung der Patienten umsetzen lassen.

3.8 Wer profitiert?

Der „Profit“ drückt sich in den Number needed to treat (NNT) aus. Diese beschreibt, wie viele Personen behandelt werden müssen, um eine einzige Person von einer Behandlung profitieren zu lassen. Die Rechnung ist einfach: Teilt man die Zahl 100 durch den Prozentsatz derer, die absolut von der entsprechenden Maßnahme profitiert haben, erhält man die Anzahl von Personen, die behandelt werden müssen, um einer einzigen Person einen Nutzen zu bringen (1/ARR).

Eine Empfehlung, ab welcher NNT der Patient behandelt werden muss, gibt es nicht. Doch im Gespräch mit dem Patienten wird dieser rasch entscheiden können, ob er eine Therapie beginnen will, von der einer von 10 oder einer von 100 Behandelten profitiert.

Die Number needed to harm (NNH) beschreibt die für die Therapieentscheidung interessante Häufigkeit von Nebenwirkungen einer Maßnahme. Eine wichtige Information, wenn der Arzt den Patienten umfassend beraten will.

Die Zusammenschau beider Wahrscheinlichkeiten wird die Entscheidung des Patienten beeinflussen: Die Chance, dass eine Behandlung nutzen wird, muss abgewogen werden gegenüber der Gefahr, dass diese Behandlung schaden könnte.

Szenario

Der Effekt von 0,51% ist den Teilnehmern des Qualitätszirkels viel zu gering, um die Prophylaxe wie von der Klinik empfohlen zu ändern. Denn berechnet man daraus die Anzahl von Patienten, die behandelt werden müssten, um ein kardiovaskuläres Ereignis zu verhindern (NNT= 1/0,51%), ergibt sich die stolze Zahl von 196 Patienten, die Clopidogrel erhalten müssten, um ein Ereignis pro Jahr zu verhindern. 195 Patienten bekämen folglich unnötigerweise das teure Mittel – entweder weil sie trotz Clopidogrel einen

Insult oder Herzinfarkt erleiden oder auch unter ASS nicht erkranken würden. Kopfschütteln in der Gruppe.

Einmal in Fahrt, analysieren die Kollegen auch die beschriebenen Studienendpunkte. Waren die „weichen Parameter“ wie Laborwerte oder handelte es sich um „harte Parameter“ wie Behinderung oder Tod? Wurden in der Studie somit Ereignisse untersucht, die die Lebensqualität der Patienten mindern können? Dies ist hier der Fall: Apoplex und Herzinfarkt stellen in der Tat für die Patienten wichtige, das weitere Leben beeinflussende Faktoren dar. Die Kollegen beschließen, diese Endpunkte für ihre Behandlung zu akzeptieren.

3.9 Sinnvolle Ergebnisse?

Ist das, was in Studien gemessen wurde, überhaupt wichtig? Dies ist eine Frage, die sowohl für den Patienten wie auch den Arzt von grundlegender Bedeutung ist. Geht es für den Patienten um so gravierende Ereignisse wie Tod oder Invalidität (Harte Endpunkte) oder handelt es sich um die Verbesserung medizinischer Parameter wie Cholesterin oder Nierenwerte (Surrogatparameter). Um sich nicht aufs Glatteis führen zu lassen, muss selbst bei harten Endpunkten wie Tod gefragt werden, was genau bestimmt wird: Misst man den Tod aufgrund von Herz-Kreislaufkrankungen oder die Mortalität überhaupt? Dieser Unterschied ist für den Wissenschaftler möglicherweise interessant – für den Patienten jedoch nur von marginalem Interesse. Hier kann der in EbM versierte Arzt für seinen Patienten die Spreu vom Weizen trennen [15].

Szenario

Die pfiffigen Kollegen rechnen sich nun noch die Kosten der Behandlung aus. Clopidogrel ist 81-mal teurer als ASS. Ein verhindertes kardiovaskuläres Ereignis würde also durch den Einsatz von Clopidogrel 81 x 196 = 15 876-mal teurer zu

Buche schlagen als die Sekundärprävention mit ASS.

Unter diesen Voraussetzungen entscheidet der Qualitätszirkel, Clopidogrel nur noch bei Patienten zu verwenden, die auf eine Prophylaxe mit ASS allergisch reagieren.

3.10 Nachhaltige Zufriedenheit mit der eigenen Arbeit – kein Traum

Immer wieder verunsichern Zweifel den Arzt an der Richtigkeit seiner Behandlungsstrategie, erfährt er doch regelmäßig, dass sogar international gültige Leitlinien von den Interessen Dritter beeinflusst werden [16]. Doch der Gefahr, durch Fehlinformationen gelenkt zu werden und willfähriges Objekt fremder Interessen zu sein, lässt sich einfach und kompetent begegnen: Durch valide Studien erhärtete und aussagekräftige Belege können die ärztliche Beratung auf ein medizinisch hohes und unabhängiges Niveau heben. Ein Niveau, das den Arzt frei von pekuniären Eigen- und Fremdinteressen macht und in Kombination mit seiner Erfahrung und dem Patientenwunsch seine Behandlung auf eine möglichst sichere und erfolgreiche Basis stellt. Dergestalt auf der „sicheren Seite“ stehend, kann die Freude am ärztlichen Handeln wieder ungetrübt genossen werden.

Szenario

Frau Stern ist zufrieden, einen Weg gefunden zu haben, das Medikamentenbudget zu schonen und ihre Patientin nachweislich sinnvoll und richtig zu behandeln.

Literatur

- [1] Ollenschläger G et al., Über die Umsetzung „evidenzbasierter Medizin“ in den Alltag einer Allgemeinpraxis. Tagl Prax. (2003), 44, 817–828

- [2] Wächtler H, Begegnung mit Evidenz-basierter Medizin aus einer Landpraxis heraus. Schleswig Holstein Arztebl. (2002), 62–63
- [3] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). AWMF online. <http://www.awmf-online.de> (08.01.07)
- [4] Ollenschläger G, Kirchner H, Fiene M, Leitlinien in der Medizin – scheitern sie an der praktischen Umsetzung? Internist (Berl). (2001), 42, 473–483
- [5] Raspe H (2000) Möglichkeiten und Gefährdungen der evidenzbasierten Medizin in Deutschland. In: Kunz R et al., Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis, 50–61. Dt. Ärzte-Verlag, Köln
- [6] Kaiser T et al., Sind die Aussagen medizinischer Werbeprospekte korrekt? Arznei-Telegramm. (2004), 35, 21–23
- [7] Boyd CM et al., Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA. (2005), 294, 716–724
- [8] Politics. Prescrire Int. (2005), 14, 192
- [9] Korzilius H, Arzneimittelwerbung: Darf's ein bisschen mehr sein? Dtsch Arztebl. (1998), 95, A–2281
- [10] Sadowski EM et al., Evaluation komplexer Interventionen: Implementierung von AR-
RIBA-Herz?, einer Beratungsstrategie für die Herz-Kreislaufprävention. ZFA. (2005), 81, 429–434
- [11] Zeh G et al., (2000) Integration der EbM in die studentische Ausbildung – Beispiel Ulm. In: Kunz R et al., Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis, 376–379. Dt. Ärzte-Verlag, Köln
- [12] Diener HC et al., Aspirin and clopidogrel compared with clopidogrel alone after recent ischaemic stroke or transient ischaemic attack in high-risk patients (MATCH): randomised, double-blind, placebo-controlled trial. Lancet. (2004), 364, 331–337
- [13] Bhatt DL et al., Clopidogrel and Aspirin versus Aspirin Alone for the Prevention of Atherothrombotic Events. N Engl J Med. (2006), 1706–1717
- [14] CAPRIE Steering Committee, A randomised, blinded trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). Lancet. (1996), 348, 1329–1339
- [15] Donner-Banzhoff N et al., Der Evidenzbasierte Praktiker. Ein Beitrag zum hausärztlichen Informationsmanagement. ZFA. (2003), 79, 501–509
- [16] Taylor R, Giles J, Cash interests taint drug advice. Nature. (2005), 437, 1070–1071

4 Welche Evidenz braucht der Arzt?

Thomas Kühlein, Johannes Forster

Um die Frage beantworten zu können, welche Evidenz der Arzt braucht, kommt man nicht umhin, darüber zu sprechen, was Evidenz ist und wozu man sie braucht.

Evidenz (evidence) ist alles, was uns glauben macht, dass etwas der Fall ist. Man unterscheidet zwischen interner und externer Evidenz. Die interne Evidenz ist die Summe der Befunde, Erfahrungen (Expertise), Meinungen und Vorstellungen, die Arzt und Patient von jeweils eigener Seite in eine Begegnung mit einbringen. Unter externer Evidenz werden Befunde aus der wissenschaftlichen Literatur verstanden.

Zur Ausübung von Evidenzbasierter Medizin braucht der Arzt neben seiner klinischen Expertise diejenige externe Evidenz, die es ihm ermöglicht, mit seinem Patienten die beste Entscheidung zu treffen. Eine brauchbare Studie muss gewissen Qualitätsansprüchen genügen. Die Patienten der Studie müssen dem individuellen Patienten möglichst ähnlich sein. Die Studie muss nach für den Patienten bedeutsamen Endpunkten fragen. Die Effekte der in Frage stehenden Intervention sollten in absoluten, nicht nur in relativen Verhältniszahlen dargestellt werden. Die Entscheidung über die Relevanz der Studienergebnisse ist unter Berücksichtigung der beiderseitigen internen Evidenzen und Werte schließlich Sache des individuellen Patienten und seines Arztes.

4.1 Es gibt mehr als eine Evidenz

Der gemeinte Begriff der Evidenz (evidence) stammt aus dem Englischen und deckt sich

nicht mit dem gleich lautenden deutschen Begriff. Evidenz ist im Englischen etwa gleich bedeutend mit „Beweis“. Aber nicht der unumstößliche Beweis (proof), wie ihn die Mathematik, die Logik oder die Naturwissenschaften liefern, ist gemeint, sondern der Beweis, wie er zum Beispiel bei Gericht vorgebracht wird. Auch dort führt erst die Summe der Beweise und Gegenbeweise zu einem Urteil. So verstanden bedeutet Evidenz also eher so etwas wie Zeugnis oder (rechnerischer) „Beleg“. Ein empirisch fundiertes Argument dafür oder dagegen, dass sich etwas so oder so verhält. Evidenz führt nicht zu Wahrheiten, sondern zur Annäherung an Wahrheiten aufgrund von Wahrscheinlichkeiten. Evidenz kann mehr oder weniger glaubhaft sein und mehr oder weniger der Wahrheit oder Wirklichkeit entsprechen. So gesehen ist Evidenz also das, was der Mensch für seine Urteile über die Welt und sein Handeln in der Welt als Beweis und Begründung anführt. Medizin war somit schon immer evidenzbasiert.

Was also ist neu an der Evidenzbasierten Medizin? Dreierlei:

- ▲ Erstens das explizite Suchen und Verwenden der besten externen Evidenz,
- ▲ zweitens der explizite Abgleich dieser Evidenz mit dem Wissen des Arztes über den Patienten (interne Evidenz/Expertise) und
- ▲ drittens die explizite Aufforderung an den Patienten, seine interne Evidenz und seine Werte in den Entscheidungsprozess mit einzubringen.

Externe und interne Evidenz können ineinander übergehen. Externe Evidenz lässt

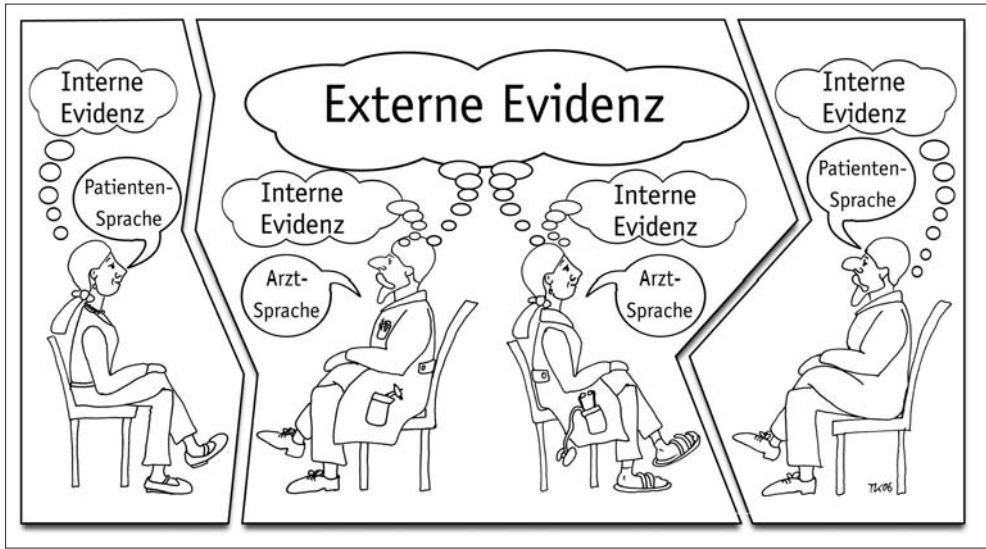


Abb. 4.1: Die Basis der gemeinsamen externen Evidenz kann Beliebigkeit verhindern, interne Evidenz dagegen ermöglicht Individualität in der Medizin.

sich je nach Beweiskraft in verschiedene Stärkegrade einteilen. Dabei bilden den niedrigsten Evidenzgrad die Meinungen von Experten und die klinische Erfahrung aus Fallberichten. Den höchsten Grad bildet die Evidenz aus Meta-Analysen randomisierter kontrollierter Studien. Der niedrigste Evidenzgrad überschneidet sich also mit der internen Evidenz des Experten. Es muss klar sein, dass die Expertenmeinung nicht deshalb so niedrig eingestuft wird, weil die Experten nicht die richtige Meinung haben könnten, sondern weil die Verlässlichkeit der Expertenmeinung aufgrund ihrer Varianz gering ist. Eine lange bekannte externe Evidenz wird sich im Lauf der Zeit und durch die Erfahrung mit den auf ihr gegründeten Entscheidungen zunehmend internalisieren.

Auch der Patient besitzt interne Evidenz. Sie bestimmt entscheidend mit, ob die vorgenannten Arten der Evidenz eine Relevanz für ihn haben.

4.2 Validität externer Evidenz

Man unterscheidet zwischen interner und externer Validität von Evidenz. Interne Validität bezieht sich auf die Qualität der Studie. Externe Validität dagegen bezieht sich auf den Wert der Studie im individuellen Kontext. Der Arzt braucht Evidenz mit hoher interner und externer Validität.

Wie schon gesagt, gibt es Evidenz unterschiedlicher Güte, die in Evidenzgraden beschrieben wird. Die Bewertung der qualitativen Unterschiede innerhalb der einzelnen Evidenzgrade ist Inhalt des so genannten „Critical Appraisal“ (kritische Beurteilung). Das Critical Appraisal beurteilt also die interne Validität einer Studie. Dies ist eine Teildisziplin der Evidenzbasierten Medizin. Nicht alle kontrollierten randomisierten Doppelblind-Studien sind automatisch gut und bedeutsam. Jede Studie hat diskutierbare Stärken und Schwächen, so dass die Validität von Studienergebnissen nicht allein vom Studiendesign, also vom Evidenzgrad her, bestimmt wird (s. Kap. 35).

4.3 Der Arzt braucht externe Evidenz, die eine konkrete Frage aus der Situation eines individuellen Patienten beantworten hilft

Die Antwort auf die Frage, welche Evidenz der Arzt braucht, lässt sich verkürzen auf die Aussagen, dass er erstens Evidenz benötigt, die an Patienten gewonnen wurde, die dem vor ihm sitzenden individuellen Patienten möglichst ähnlich sind. Mit anderen Worten, der Arzt braucht Evidenz mit aus seiner Sicht hoher externer Validität. Zweitens braucht er Evidenz aus Studien, die eine für den individuellen Patienten bedeutsame Frage beantworten. Was nützt die Studie mit der größtmöglichen Qualität, wenn sie keine relevante Frage beantwortet?

Ein entscheidendes Auswahlkriterium für die benötigte Evidenz ist also, ob die untersuchten Patienten dem individuellen Patienten ähnlich sind. Welche Bedeutung hat wohl eine an 45852 Chinesen gewonnene Evidenz [1] für die fränkische Bäuerin mit frischem Herzinfarkt? Auch sind mit großer Sicherheit die Patienten einer Universitätsklinik ganz anders gartet als die, denen der Landarzt im Wochenenddienst begegnet.

Die Frage nach der Relevanz dreht sich um zwei Dinge. Erstens, ob die Studie überhaupt eine für den Patienten relevante Frage stellt. Und zweitens, ob der gefundene Therapieeffekt eine für den Patienten relevante Größe hat.

Ein effektives Entscheidungskriterium für ersteres ist die Unterscheidung zwischen DOE und POEM [2]. DOE bedeutet „Disease oriented evidence“. Es handelt sich entweder nicht um klinische Forschung, das heißt die Studie beschäftigt sich beispielsweise mit Versuchstieren oder Zellmembranen. Oder die Studie fragt nach Endpunkten, die zwar das Krankheitsgeschehen wahrscheinlich beeinflussen, vom Patienten selbst aber nicht wahrgenommen werden können. Sie besit-

zen keine subjektive Bedeutung für ihn. Solche Endpunkte nennt man auch Surrogatparameter. Ein Beispiel wäre die Konzentration von Hausstaubmilben im Bett nach milbendichtem Überzug der Matratze. Ein milbenreduzierender Effekt lässt sich nachweisen [3].

DOEs untersuchen die Richtigkeit pathophysiologischer Hypothesen. DOEs sind wichtig, um neue Wege zu finden, Patienten zu helfen. DOEs sind wichtig für Forscher.

POEM bedeutet „Patient oriented evidence that matters“. Evidenz also, die aus Studien stammt, die sich direkt mit Patienten beschäftigen und die für diese direkt erfahrbare und bedeutsame Fragen beantworten. Im Falle der Allergien gegen Hausstaubmilben wäre dies also die Frage, ob milbendichte Überzüge die Symptome der gequälten Patienten reduzieren können. Dies scheint, so elegant der Therapieansatz auch ist, leider nicht wesentlich der Fall zu sein [2; 4; 5].

POEMs überprüfen die tatsächliche Auswirkung von über DOEs gewonnen pathophysiologischen Vorstellungen und daraus gewonnenen Therapieansätzen in der Wirklichkeit der Patienten. Der klinisch tätige Arzt braucht POEMs, nicht DOEs.

4.4 Der Arzt braucht schnell verfügbare Evidenz, der er vertrauen kann

Alle Evidenz ist Menschenwerk. Medizin und medizinische Forschung sind ein kulturelles Phänomen und unterliegen damit Einflüssen wie Zeitgeist, Wünschen, Hoffnungen, politischen und finanziellen Interessen. Diese beeinflussen die Forschung und die Medizin von der Wahl des Forschungsfeldes bis zur Interpretation der Ergebnisse.

Der Arzt braucht Evidenz, die schnell verfügbar ist. Dafür braucht er verlässliche Quellen, die für ihn Evidenz lesen, bewerten

und zusammenfassen. Um der Evidenz vertrauen zu können, sollte er sie aber am besten selbst gelesen und beurteilt haben.

Viele Fragen lassen sich schnell über vorbereitete Aufarbeitungen, wie zum Beispiel systematische Übersichtsartikel (reviews), wie sie die Cochrane Library bietet, evidenzbasierte Leitlinien, das Kompendium Evidenzbasierte Medizin [6] (Clinical Evidence concise) oder InfoPoems [7], aber auch über viele andere Wege beantworten. Der Weg zum Vertrauen in die externe Evidenz führt jedoch am besten über das eigene Urteil mit Hilfe grundsätzlicher Kenntnisse der kritischen Bewertung, des Critical Appraisal (s. Kap. 9–18). Auch Übersichtsartikel und Leitlinien können und sollten einem Critical Appraisal unterzogen werden. Die Transparenz ihrer Entstehung spielt dabei eine entscheidende Rolle.

4.5 Verschiedene Ärzte benötigen verschiedene Evidenz

Verschiedene Ärzte betreuen verschiedene Patienten mit verschiedenen Krankheiten und haben dabei verschiedene Aufgaben. Wir wollen im Weiteren versuchen die Unterschiede zwischen Generalisten und Spezialisten heraus zu arbeiten, auch wenn zwischen diesen keine klare Grenze besteht.

4.5.1 Verschiedene Evidenz für verschiedene Patienten

Generalist und Spezialist betreuen in der Regel völlig unterschiedliche Patienten. Es beginnt damit, dass der Patient des Generalisten selbstständig entscheidet, welches Symptom ihn derartig beunruhigt, dass er den Arzt zu Rate ziehen möchte. Der Spezialist dagegen bekommt häufig nur die von Kollegen vorher ausgewählten Patienten zu Gesicht. Patienten sind, bevor sie zum Arzt gehen, kei-

neswegs unberaten oder gar unbehandelt. Soziokulturelle Eigenheiten, eigene Vorstellungen und Erfahrungen und nicht zuletzt die Organisation des Gesundheitswesens beeinflussen die Schwelle, den Arzt aufzusuchen. Die Patienten berichten in ihren eigenen Worten von dem, was sie zum Arzt führt. Der Generalist sieht Patienten mit einer großen Varianz von Beschwerden. Die Beschwerden variieren nicht nur bezüglich ihrer Art, sondern auch bezüglich ihres Schweregrades und des Zeitpunktes im natürlichen Verlauf der Erkrankung. Das Verhältnis des Patienten zu seinem Hausarzt ist paritätischer als das zu dessen spezialisierten Kollegen. Vorstellungen und Wünsche der Patienten, also deren innere Evidenz, haben einen deutlich höheren Stellenwert. Der Generalist braucht also eine bessere Kenntnis der inneren Evidenz des Patienten als der Spezialist. Ein Patient im Herzkatheterlabor wird sicher weniger mit seinem Arzt über die richtige Behandlung diskutieren als der hausärztliche Patient, dem gerade geraten wurde, ein drittes Antihypertensivum einzunehmen.

Die Freiheit der Wahl des Arztes schränkt sich faktisch mit dessen Spezialisierung zunehmend ein. Das bedeutet auch, dass das Verhältnis des Generalisten zu seinem Patienten persönlicher ist. Hausärzte betreuen ihre Patienten nicht nur über längere Zeiträume, sondern sie begegnen ihnen auch öfter und kümmern sich nacheinander und oft auch gleichzeitig um sehr unterschiedliche Probleme der Patienten. Hausärzte betreuen in vielen Fällen gleichzeitig die Eltern, Ehegatten und Kinder ihrer Patienten. Alle diese Dinge haben Einfluss auf die zu treffenden Entscheidungen. Alle diese Dinge setzen externe Evidenz in individuellen Kontext.

Je spezialisierter der Arzt, desto geringer wird die Varianz der Krankheiten seiner Patienten, während die Varianz der Lebenskontexte seiner Patienten erhalten bleibt. Bezüglich der Krankheiten ist durch die vorgehalteten Generalisten eine erhebliche

Auswahl und Sortierung eingetreten. Die Patienten sind längst durch die Vorstellungen und Erklärungen der überweisenden Ärzte in ihren Ansichten und Erwartungen beeinflusst. Ein Kardiologe im Herzkatheterlabor wird sich nicht um die Schlafstörungen seiner Patienten kümmern, und weder der Patient noch der überweisende Arzt würden diese als dessen Aufgabe betrachten. Dass ein die Brust einschnürendes Gefühl in Verbindung mit den Schlafstörungen statt als Angina pectoris als Zeichen einer Depression gesehen werden könnte, wird ihm nicht in den Sinn kommen, selbst wenn er es weiß. Der Spezialist fragt stattdessen nach Nikotinstatus, Cholesterinwert, Bluthochdruck, Diabetes und Familienanamnese für koronare Herzkrankheit. Weder überweisender Arzt noch Patient würden anderes erwarten. Überspitzt formuliert trägt der Patient sein Herz zum Kardiologen, damit der sich darum kümmere. Hält der Kardiologe das Herz des Patienten für gesund, wird er dies dem Patienten und seinem Hausarzt mitteilen und ihn nach Hause schicken, unbekümmert darum, ob sein Problem weiterhin besteht.

Der Experte arbeitet in einer Art Labor-situation. Ein großer Teil der externen Evidenz stammt aus Studien, die in spezialisiertem Setting durchgeführt wurden. Der Generalist kann diese Evidenz nur begrenzt auf seine Patienten übertragen. Folglich ist es für den Generalisten erstens sinnvoll, die Herkunft der Evidenz zu überprüfen. Zweitens sollte er sich an hausärztlich und nicht an fachärztlich erstellten Leitlinien orientieren.

4.5.2 Verschiedene Evidenz für verschiedene Krankheiten

Im Englischen unterscheidet man zwischen den Begriffen „illness“ und „disease“. Vielleicht kann man diese Begriffe am ehesten mit Kranksein und Krankheit übersetzen.

Dabei ist Kranksein (illness) der weiter gefasste Begriff. Er umfasst die Wahrnehmung einer körperlichen, geistigen oder seelischen Störung und alles, was aus ihr folgt. Man kann Kranksein (illness) als das bezeichnen, was den Patienten zum Arzt führt und Krankheit (disease) als das, womit er wieder nach Hause geht [8; 9]. Die Krankheit ist somit ein Konstrukt unserer medizinischen Vorstellungen und Lehrgebäude [10]. Kranksein ist basal ein natürliches Phänomen, Krankheit in Form einer definierten Diagnose dagegen ein kulturelles. Es ist das Ergebnis ärztlicher Arbeit, aus dem unendlichen Spektrum der möglichen Symptome definierte Krankheitsentitäten zu destillieren und ihnen in Form von Diagnosen Namen zu geben. Die Benennung eines individuellen Krankseins mit dem Namen einer Krankheit stellt eine Verallgemeinerung dar. Die zugrunde liegenden Krankheitskonzepte und Vorstellungen unterliegen einem stetigen Wandel [11]. Medizin ist in weiten Teilen die Behandlung des Krankseins eines individuellen Patienten auf der Basis der Erfahrung mit anderen ähnlichen Patienten. Die Ähnlichkeit wird über den Namen der Krankheit definiert. Der Generalist arbeitet dabei näher am Kranksein (illness), der Spezialist näher an der Krankheit (disease). Vieles, was der Generalist sieht, bleibt im Unschärfer und Symptomhaften. Vieles davon ist harmloser Natur in dem Sinne, dass es auch ohne Therapie wieder verschwinden würde. Vieles davon ist chronisch und nicht zu ändern. Vieles gehört ganz normal zu den verschiedenen Lebensphasen der Patienten. Die wenigsten der von Patienten beschriebenen Symptome lassen sich klaren Diagnosen zuordnen. Ärzte wollen heilen und Patienten wollen geheilt werden. Oft ist Heilung aber nicht möglich und so wandelt sich die kurative Tätigkeit des Arztes (cure) in eine karitative (care), betreuende, begleitende. Auch hierin unterscheiden sich die Krankheitsfälle der Spezialisten und der Generalisten.

Die externe Evidenz der klinischen Studien bezieht sich in der Regel auf Krankheiten und nicht auf das Kranksein. Die klassische quantitative Forschung benötigt die Verallgemeinerung, die durch die Definition einer Krankheit entsteht, im Sinne einer *conditio sine qua non*. Aufgabe der ärztlichen Tätigkeit ist der Schritt aus der Verallgemeinerung der Forschung wieder zurück zum Individuum. Dafür benötigt der Arzt seine eigene innere Evidenz, die des Patienten und vor allem Kommunikationsfähigkeiten.

Über das Kranksein, beziehungsweise die individuelle Bewertung von Kranksein und Therapieoptionen gibt es ebenfalls externe Evidenz in Form so genannter qualitativer Forschung (s. Kap. 14). Diese findet zunehmend ihren Platz in der hausärztlichen Literatur.

4.5.3 Verschiedene Evidenz für verschiedene Aufgaben

Nicht nur Patienten sind Individuen. Ärzte sind keine genormten Diagnose- und Therapiemaschinen, die streng rational aufgrund von Forschungsergebnissen zu den immer gleichen Schlussfolgerungen kämen. Ein Urteil über dieselbe Evidenz kann individuell sehr unterschiedlich ausfallen. Auch ist ein Psychiater nicht zufällig ein Psychiater. Er ist es geworden aufgrund ihm eigener Charaktereigenschaften und Vorbilder, aufgrund von Gegebenheiten und Chancen. Er erfährt im Laufe seiner Ausbildung und Arbeit eine vollkommen andere Sozialisation als beispielsweise ein Gefäßchirurg. Der Chirurg wird tendenziell versuchen, die Probleme seiner Patienten mit dem Skalpell zu lösen, der Internist medikamentös und der Psychotherapeut eben psychotherapeutisch. Jeder von ihnen sucht und benötigt eine andere Evidenz, um sein Handeln zu begründen.

Generalisten sind wiederum ganz eigene Typen. Vermutlich sind sie untereinander

deutlich heterogener als die meisten Gruppen von Spezialisten. Sie haben meist sehr unterschiedliche Ausbildungen erhalten. Sie stehen alternativen Heilmethoden oft weit näher und der wissenschaftlichen Datenlage ferner als ihre spezialisierten Kollegen. In Deutschland scheinen die meisten Hausärzte ihre Entscheidungen in der Regel auf ihre interne Evidenz in Form ihrer persönlichen Erfahrung zu gründen, wohingegen Fachärzte oft tief in ihren pathophysiologischen Vorstellungen verhaftet sind. Hausärzte beziehen ihre Evidenz aus anderen Quellen als Spezialisten. Noch zu wenige lesen Originalpublikationen und haben gelernt diese selbst zu bewerten. Sie verlassen sich dagegen häufig auf deren Interpretation und Zusammenfassung durch Spezialisten, die mit der hausärztlichen Situation wenig vertraut sind.

Große Teile des Krankseins haben weder einen konkreten Namen, noch bedürfen sie einer spezifischen Therapie, noch gibt es irgendwelche externe Evidenz für ihre Behandlung. „Managing Uncertainty“ ist ein stehender Ausdruck der allgemeinmedizinischen Theorie und ein alltägliches Problem in der Praxis.

Bonmot: In einer berufspolitischen Diskussion äußerte ein fachärztlich tätiger Internist: „Aber ich kann doch nicht Verantwortung für etwas übernehmen, von dem ich keine Ahnung habe.“ Darauf lachte ein hausärztlicher Kollege freundlich und sagte: „Siehst du, das ist der Unterschied zwischen Hausarzt und Facharzt – wir machen den ganzen Tag nichts anderes.“

Der Spezialist benötigt Evidenz für sein Handeln bei definierten Krankheiten. Der Generalist benötigt Evidenz oft für das Schwierigste, was der Arzt tun kann: das begründete Nicht-Handeln [12]. Er braucht an vielen Stellen Evidenz – nicht um Krankheiten zu behandeln, sondern um sie mit größtmög-

cher Sicherheit auszuschließen oder ein abwartendes Offenlassen zu rechtfertigen. Nur in Kenntnis von Wahrscheinlichkeiten und absoluten Effektgrößen ist der Arzt in der Lage, auf die für den Patienten am wenigsten relevanten diagnostischen, präventiven und therapeutischen Maßnahmen zu verzichten.

4.5.4 Übergang von Evidenz zwischen den Versorgungsbereichen

Michael Balint nannte die Beziehung zwischen Hausarzt und Facharzt schon 1964 die „Verewigung des Lehrer-Schüler-Verhältnisses“ [13]. Daran hat sich in Deutschland erst wenig geändert. Die institutionalisierte Medizin in Deutschland wird als Schulmedizin bezeichnet. Die Information (Evidenz) fließt entlang eines hierarchischen Gefälles von den Spezialisten der Universitäten und ihren Kliniken, Labors und Instituten über die Fachärzte zu den Hausärzten, den Generalisten. Die Schulmedizin gründet ihre Entscheidungen traditionell auf pathophysiologischen Erkenntnissen. Erkenntnissen also, die aus DOEs stammen. Zu großen Teilen bezieht der Generalist seine externe Evidenz als Expertenmeinung, die aufgrund ihrer Variabilität als niedrigste Stufe externer Evidenz gesehen wird.

Demgegenüber hat sich eine blühende Kultur der so genannten Alternativmedizin entwickelt. Alternativmediziner und auch viele Schulmediziner neigen zu dem Fehler, die Ergebnisse ihres Handelns im Vergleich zum Nicht-Handeln oder zum Scheinmedikament nicht konsequent am Patienten zu überprüfen.

Evidenzbasierte Medizin ist empirische Medizin mit rationaler Methodik. Sie fragt nicht wie die rationalistische Schulmedizin, warum etwas helfen könnte, sie fragt, ob etwas hilft, mit welcher Wahrscheinlichkeit und in welchem Maße. Evidenzbasierte Medizin stützt sich im Gegensatz

zur ebenfalls empirischen, aber auch dogmatischen Alternativmedizin nicht auf die trügerische persönliche Erfahrung allein, sondern zusätzlich auf geregelte (kontrollierte) Empirie [14]. Zur Beantwortung der Frage, welche Evidenz der Arzt braucht, lässt sich also sagen: er braucht externe Evidenz aus geregelter empirischer Forschung. Am besten kommt die Evidenz deshalb aus randomisierten kontrollierten klinischen Doppelblindstudien. Nur in Kenntnis der absoluten Effektgrößen aus den richtigen Studien kann der Arzt die individuelle Relevanz einer Maßnahme für seinen Patienten abschätzen. Diese Kenntnis befreit ihn aus dem „ewigen Lehrer-Schüler-Verhältnis“.

4.6 Parameter der Relevanz

Signifikant ist nicht gleich relevant. Häufig reduziert sich die Interpretation einer klinischen Studie auf die Aussage, ein Medikament habe sich als signifikant wirksam erwiesen. Signifikanz bedeutet aber im Sprachgebrauch der Statistik nur, dass der zwischen den verschiedenen Gruppen gefundene Unterschied, also der Therapieeffekt, mit einer hohen Wahrscheinlichkeit (mindestens 95%) kein Zufall war. Weder der Begriff „signifikant“, noch der Begriff „hoch signifikant“ sagen irgendetwas über die Größe des gefundenen Effekts. Nicht Signifikanz, sondern die Größe des Therapieeffekts ist das Kriterium für seine Relevanz. Dabei ist Signifikanz natürlich eine Voraussetzung, das Therapeutikum überhaupt in Betracht zu ziehen.

Relativ ist nicht absolut. Die Mittelwertunterschiede zwischen behandelten und unbehandelten Patienten beantworten die Frage nach der Größe des Effekts der Intervention. Um die Effektgröße eindrücklich zu machen, wird sie gerne als Relative Risiko-Reduktion angegeben. So wurde jüngst in ei-

ner groß angelegten Studie bezüglich gastrointestinaler Blutungen durch Coxibe eine 77%ige Risiko-Reduktion gegenüber traditionellen Antirheumatika gefeiert [15]. Die kumulative Inzidenz gastrointestinaler Blutungen lag in dieser Studie nach einem Jahr unter traditionellen Antirheumatika bei 1,09% der Patienten, beim untersuchten Coxib dagegen bei 0,25%. Die Absolute Risiko-Reduktion lag also bei 0,84%. Dasselbe Ergebnis lässt sich also als knapp 80%ige und als ca. 0,8%ige Risiko-Reduktion darstellen. Beides sind sachlich richtige Formen derselben Evidenz. Welche Form der Evidenz braucht der Arzt? Für die Einschätzung des Therapievorteils im Bereich der Nebenwirkungen von Antirheumatika, wie auch für alle anderen Therapieeffekte, ist für den individuellen Patienten allein die Absolute Risiko-Reduktion von Interesse.

Die Angabe in relativen Maßzahlen lässt den Therapieeffekt unverhältnismäßig groß erscheinen. Für die meisten Menschen noch anschaulicher als die Absolute Risiko-Reduktion in Prozent ist die Angabe in absoluten Zahlen. Im Falle des Beispiels hätten ca. drei von 1000 der mit dem Coxib behandelten Patienten eine Blutung erlitten, gegenüber ca. elf von 1000 der mit traditionellen Antirheumatika behandelten Patienten. Ob diese Evidenz für einen konkreten Patienten eine Bedeutung hat, lässt sich nur am Einzelfall diskutieren.

Die Art der Darstellung der gefundenen Evidenz beeinflusst die subjektive Einschätzung ihrer Relevanz.

4.7 Die Indikation einer präventiven Therapie

Bei den großen Krankheitsbildern wie Diabetes mellitus, Bluthochdruck und koronare Herzkrankheit ist medizinische Tätigkeit weitgehend präventiv. Für präventive Therapieformen ist genügend Zeit, sich grundsätz-

lich mit der Evidenzbasis auseinander zu setzen, diese zum Beispiel in Qualitätszirkeln zu diskutieren und sie dann mit Hilfe von Qualitätsmanagement-Strukturen umzusetzen, soweit es individuell machbar und sinnvoll erscheint [16; 17]. Gerade bei Fragen der Prävention versagt die interne Evidenz als Grundlage von Entscheidungen fast vollständig. Man kann keine eigenen Erfahrungen mit seltenen Ereignissen machen, die durch eine präventive Therapie **nicht** eintreten. Der Arzt ist auf externe Evidenz angewiesen.

Nehmen wir an, ein Patient steht vor der Frage, ob er ein Antihypertensivum einnehmen soll oder nicht. Bluthochdruck ist im Wesentlichen ein Risikofaktor für Herzinfarkt und Schlaganfall. Um die Relevanz eines bestimmten Antihypertensivums für einen konkreten Patienten abschätzen zu können, sind zwei Informationen essenziell. Erstens braucht der Arzt Evidenz, um das individuelle Risiko des Patienten für Herzinfarkt und Schlaganfall abschätzen zu können. Zweitens benötigt er Evidenz, um abschätzen zu können, um wieviel dieses Risiko durch das in Frage stehende Antihypertensivum gesenkt werden kann. Erst in der ungefähren Kenntnis dieser Zahlen lässt sich die Relevanz der Medikation für den individuellen Patienten beurteilen. Bei den großen Volkskrankheiten wie Herzinfarkt und Schlaganfall und ihren Risikofaktoren Alter, Nikotinstatus, Hypercholesterinämie, Bluthochdruck, positive Familienanamnese und Diabetes lassen sich solche Fragestellungen über Risikokalkulatoren auch im klinischen Alltag leicht bewerkstelligen. Ist das individuelle Risiko klein, kann auch das wirksamste Medikament keine große Risiko-Reduktion bewirken. Ist das Risiko hoch, genügt meist nicht eine einzelne Maßnahme, um es auf ein erträgliches Maß zu reduzieren. Dazwischen liegt ein Kontinuum.

Einer der stärksten Risikofaktoren für viele Krankheiten ist das Alter. Präventive Maßnahmen für 95-Jährige erscheinen spontan

eher sinnlos. Aber wo liegt die Grenze? Die Grenze ist abhängig von der Komorbidität und natürlich von den persönlichen Sicherheitsbedürfnissen und Lebenseinstellungen des Patienten. Alle diese Variablen bilden Kontinua von Risiken, die in vereinfachender Weise oft in eine Dichotomie zwischen normal und abnormal überführt werden. Die dazugehörigen Grenzwerte lassen sich für bestimmte Szenarien optimieren, sind aber nie universal gültig. Dies gilt speziell für Krankheiten, die eigentlich Risikofaktoren sind, wie etwa Hypercholesterinämie, Bluthochdruck und Osteoporose.

Auf der anderen Seite der Relevanzeinschätzung stehen dem zu erwartenden Nutzen potenzielle Nebenwirkungen gegenüber. Erst die realistische Abschätzung des Gesamtrisikos (zu erkranken und Nebenwirkungen zu erleiden) erlaubt in Kombination mit der Kenntnis der Therapieeffekte auf den verschiedenen Risikohöhen die Einschätzung der Relevanz einer Maßnahme. Für alles braucht der Arzt externe Evidenz. Aber nicht die Evidenz selbst entscheidet über ihre eigene Relevanz. Der Arzt entscheidet über sie – für und mit seinem Patienten. Ärztliche Tätigkeit in Form von EbM macht nur Sinn im individuellen lebensgeschichtlichen Kontext des Patienten [18].

4.8 Die Akzeptanz einer Prävention

Auch wenn viele Patienten im Sprechzimmer des Arztes noch ergeben nicken und mit Rezept und neuem Medikamentenplan die Praxis oder das Krankenhaus verlassen, so bedeutet das noch längst nicht, dass sie die Behandlung auch durchführen. Zu oft bleibt die interne Evidenz des Patienten bei der Therapieentscheidung unberücksichtigt. Die Folge wird als „schlechte Compliance“ (Gehorsam) bezeichnet, sollte aber besser „fehlende Concordance“ (Übereinstimmung) heißen. Der Arzt braucht also neben der ex-

ternen Evidenz und der eigenen internen Evidenz und Expertise auch noch dringend Kenntnis von der inneren Evidenz des Patienten [19; 20].

4.9 Probleme im Alltag

Man muss sich klar sein, dass die Einschätzung der Relevanz einer Maßnahme eine Zumutung für Arzt und Patient darstellt. Die Sichtweise: „Nimm das Antihypertensivum, dann bekommst du keinen Herzinfarkt!“ beispielsweise vereinfacht zwar die Entscheidung für beide Seiten erheblich, entspricht aber nicht der Wirklichkeit. In welchem Umfang man bereit ist, sich mit rationalen Erwägungen heranzuschlagen, bevor man sich entscheidet, ist individuell unterschiedlich. Gerd Gigerenzer und Reinhard Selten nennen es die „Stopping Rule“ [21]: Wann hören wir auf, nach weiterer Evidenz zu suchen, bevor wir unsere Entscheidung treffen? Wie viel Evidenz braucht der Arzt?

Sie stellen die Grundannahme infrage, dass wir die Dinge des Lebens rational entscheiden würden. Wir gehen davon aus, wir hätten für unsere Entscheidungen alle nötigen Informationen (Evidenz) zur Verfügung, hätten endlos Zeit für die Entscheidung und unbegrenzte Rechenkapazität, um alle Eventualitäten in unsere Entscheidung mit einzuwoben. Dies ist in der Wirklichkeit fast nie der Fall und selbst wenn, dann wäre es unökonomisch. Entscheidungen fallen deshalb meist heuristisch: Heuristiken werden dabei verstanden als Entscheidungsmechanismen oder Entscheidungsregeln, die die Gesetze der Vernunft und der Wahrscheinlichkeit außer Kraft zu setzen vermögen [22]. Solche Heuristiken können heißen „Tu, was die anderen tun“, „Tu, was die Professoren sagen“, „Tu was dein Arzt dir rät“, „Tu was du schon immer getan hast und was bisher immer zum Ziel geführt hat“ und so weiter. Das Ergebnis ist ein Teil dessen, was Michael Balint

die „Zersplitterung der Verantwortung“ genannt hat. Wenn wir uns kein eigenes Urteil gebildet haben, sondern uns stattdessen nur auf Vorbilder und Üblichkeiten berufen, übernehmen wir auch nicht die Verantwortung für unser Handeln. Wir verschieben die Verantwortung auf eine andere Ebene. Der Arzt braucht deshalb externe Evidenz, um heuristisch gefällte Entscheidungen zu überprüfen und gegebenenfalls für die Zukunft zu korrigieren. Er braucht externe Evidenz, um seine Entscheidungen sicher begründen zu können und damit die Verantwortung für sie zu übernehmen.

Viele falsch getroffene Entscheidungen führen scheinbar sogar zum Erfolg. So gibt es genügend Belege, dass die Verordnung von Antibiotika bei grippalen Infekten zu keinerlei Therapievorteil führt. Dennoch werden die Patienten fast immer gesund, wenn sie Antibiotika nehmen. Die Philosophie nennt diesen Denkfehler „post hoc non est propter hoc“ – „Nach etwas ist nicht wegen etwas“. Mit anderen Worten: die Patienten werden trotz, nicht wegen der Antibiose gesund. Obwohl die meisten Ärzte das wissen, verordnen sie immer wieder Antibiotika in Situationen, an denen sie nachweislich nutzlos oder von geringem Nutzen sind. Sie handeln dabei aufgrund falscher Heuristiken, Reflexen und unreflektierter innerer Evidenz. Um dies unterlassen zu können, müssten sie die externe Evidenz besitzen, sie selbst gelesen, selbst beurteilt und – vor allem – internalisiert haben.

4.10 Leitlinien im Alltag

Eine „Anweisung von oben“ in Form einer Leitlinie genügt offensichtlich nicht, um die eingefahrenen Geleise der Gewohnheit und Üblichkeit zu überwinden. Der Arzt braucht zwar externe Evidenz, um Dinge anders zu machen als bisher und anders zu machen als die Anderen. Aber erst das Entstehen neuer

interner Evidenz bestätigt ihm auch gefühlsmäßig, auf dem richtigen Weg zu sein.

Es ist eines der erklärten Ziele Evidenzbasierter Medizin, neue wissenschaftliche Erkenntnisse schneller als über den traditionellen Weg hierarchischer Strukturen zum Patienten zu bringen. Die absolute Forderung, neue wissenschaftliche Erkenntnisse in der Patientenversorgung umzusetzen, führt zu einer berechtigten Abwehrhaltung, solange nicht zwischen individueller und epidemiologischer Relevanz unterschieden wird. Therapieeffekte, die auf die Gesamtbevölkerung gerechnet eindrucksvoll erscheinen oder ökonomisch relevant sind, müssen es für den individuellen Patienten noch längst nicht sein. Eine Leitlinie, auch wenn sie evidenzbasiert ist, ist kein Gesetz. Sie entlässt den Arzt weder aus der individuellen Entscheidung noch aus der Verantwortung für sie. Evidenzbasierte Medizin ist nicht Evidence based health care. Sie ist nicht das neue Richtschwert der Schulmedizin. Externe Evidenz ist daher nicht das Ende der Individualität und Diskussion, sondern ihre Voraussetzung, ihre Basis und ihr Beginn. In diesem Sinne: Wo die Evidenz aufhört, beginnt die Medizin.

4.11 Zusammenfassung

Evidenz bedeutet „Beweis“ im juristischen Sinne. Erst die Summe der Beweise und Gegenbeweise ermöglicht ein Urteil. Ärzte und Patienten haben schon immer ihre innere Evidenz benutzt, um zu Entscheidungen zu kommen. Evidenzbasierte Medizin nutzt zusätzlich so genannte externe Evidenz aus klinischer Forschung und macht explizit Gebrauch von allen genannten Formen der Evidenz.

Evidenzbasierte Medizin ist individuelle Medizin. Der Arzt braucht daher Evidenz aus Studien, die sich auf für den Patienten erfahrbare und bedeutsame Endpunkte stützen (POEMs). Studien liefern Wahrscheinlichkei-

ten, nicht Wahrheiten. Der Arzt braucht Evidenz nicht in Form der trügerischen Dichotomie zwischen wirksam und unwirksam, sondern in absoluten Effektgrößen.

Der Arzt braucht Evidenz, der er vertrauen kann. Dazu kann er selbst Studien lesen und sie kritisch beurteilen (Critical Appraisal) oder sich auf qualitativ gute und transparent erstellte Aufarbeitungen der Evidenz stützen.

Verschiedene Ärzte brauchen verschiedene Evidenz für verschiedene Patienten und verschiedene Formen ärztlichen Handelns. Was für die Patienten des Spezialisten Bedeutung hat, muss es nicht für die Patienten des Generalisten haben. Wo die Evidenz als Wissenschaft aufhört, beginnt die Medizin, denn die Relevanz von Studienergebnissen und Leitlinien muss mit dem einzelnen Patienten erst noch erarbeitet werden.

Literatur

- [1] Chen ZM et al., Addition of clopidogrel to aspirin in 45,852 patients with acute myocardial infarction: randomised placebo-controlled trial. *Lancet*. (2005), 366, 1607–1621
- [2] Rosser WW, Slawson DC, Shaugnessy AF (2004) Information mastery: Evidence-based family medicine. Decker, London
- [3] Terreehorst I et al., Evaluation of impermeable covers for bedding in patients with allergic rhinitis. *N Engl J Med*. (2003), 349, 237–246
- [4] Gotzsche PC et al., House dust mite control measures for asthma. *Cochrane Database Syst Rev* (2004), CD001187
- [5] Sheikh A, Hurwitz B, House dust mite avoidance measures for perennial allergic rhinitis. *Cochrane Database Syst Rev*. (2001), CD001563
- [6] Ollenschläger G et al. (2005) Kompendium evidenzbasierte Medizin 2006. Clinical Evidence Concise. Huber, Bern
- [7] InfoRetriever with InfoPOEMs. The clinical awareness system. www.infopoems.com (2006 Dez 14)
- [8] Helman CG (2000) Culture, health and illness. Butterworth, London
- [9] Cassel FJ (1978) The healer's art: A new approach to the doctor-patient relationship. Penguin, Harmondsworth
- [10] Fleck L (1980) Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv. Suhrkamp, Frankfurt/Main
- [11] Kuhn TS (1976) Die Struktur wissenschaftlicher Revolutionen. Suhrkamp, Frankfurt/Main
- [12] Siebolds M, Evidenzbasierte Medizin als Modell der Entscheidungsfindung in ärztlicher Praxis. *Z Arztl Fortbild Qualitätssich* (2003), 97, 257–262
- [13] Balint M (1964) The doctor, his patient and the illness. Pitman Med. Publ., London
- [14] Tröhler U (1999) Die wissenschaftlichen Begründungen therapeutischer Entscheide – oder „Evidence-Based medicines“ – im Lauf der Geschichte. In: Eich W et al., Wissenschaftlichkeit in der Medizin. Teil III: Von der klinischen Erfahrung zur Evidence-based Medicine, 101–127. VAS, Frankfurt/Main
- [15] Farkouh ME et al., Comparison of lumiracoxib with naproxen and ibuprofen in the Therapeutic Arthritis Research and Gastrointestinal Event Trial (TARGET), cardiovascular outcomes: randomised controlled trial. *Lancet* (2004), 364, 675–684
- [16] Van Zwanenberg T, Harrison J (Eds.) (2004) Clinical Governance in primary care. Radcliff Med. Pr., Abingdon
- [17] Baker R et al. (1999) Implementing change with clinical audit. Wiley, Chichester
- [18] Greenhalgh T, Hurwitz B (Eds.) (1998) Narrative based medicine. Dialogue and discourse in clinical practice. BMJ Books, London
- [19] Stewart M et al. (2003) Patient centered medicine-transforming the clinical method. Radcliff Med. Pr., Abingdon
- [20] Tucket D et al. (1985) Meetings between experts: an approach to sharing ideas in medical consultations. Tavistock Publ., London
- [21] Gigerenzer G, Selten R (2002) Rethinking rationality. In: Gigerenzer G, Selten R, Bounded rationality – the adaptive toolbox, 1–12. MIT Pr., Cambridge
- [22] Gigerenzer G, Todd PM, The ABC research group (1999) Simple heuristics that make us smart. Oxford Univ. Pr., Oxford

5 Welche Evidenz will der Patient?

Sylvia Sanger, Bernd Quadder, Frank Brunsmann

5.1 Einfuhrung

Im landlaufigen Verstandnis vieler Patienten verheit der Begriff Evidenz „den allwissenden Arzt auf dem Weltniveau der medizinischen Forschung und Erkenntnis“. Im Behandlungszimmer hofft der Patient, dem kompetenten Arzt gegenuber zu sitzen, der in unabhangiger Beraterposition seinen Patienten einfuhlsam durch die Untiefen seiner Erkrankung geleitet. Welcher Patient mochte bei einer schwerwiegenden oder gar lebensbedrohlichen Erkrankung nicht genau aus dieser Gewissheit Vertrauen zu seinem Arzt schopfen, wobei dieser sich als wahrer Freund erweisen moge. Dies stellt die unausgesprochene Erwartung der Mehrheit der Patienten an das Selbstverstandnis und die Selbstverpflichtung des Arztes dar. Die Enttauschung durch manche Versorgungserfahrungen im Medizinbetrieb hat weder die Einforderung evidenzbasierten arztlichen Handelns noch die Erwartung des Patienten an seinen Arzt aufgehoben. Die Bereitschaft zum Vertrauen in jeden neuen Behandler und die Hoffnung auf Genesung oder zumindest eine Verbesserung der Lebensqualitat durch die Behandlung sind stets vorhanden.

Das Ansinnen der Evidenzbasierten Medizin, die beste verfugbare Erkenntnislage in der individuellen Situation des Patienten anwendbar zu machen, hat bei Patienten Erwartungen auf viele Qualitatsverbesserungen in der Krankenversorgung ausgelost. Diese Erwartungen wurden bislang erst zu einem kleinen Teil erfullt. Aus der Sicht von Betroffenen verkehrten sie sich gelegentlich sogar in ihr Gegenteil.

Die Grunde fur die zum Teil enttauschten Erwartungen an die Evidenzbasierte Medizin sind:

- ▲ Die Fragen „Woher kommt die Evidenz?“ und „Wer definiert denn eigentlich die Evidenz?“ wurden in der Vergangenheit oft nur von gesunden Experten gestellt. Die Sicht, Erfahrungen und Bewertungen der Patienten, einschlielich der organisierten Selbsthilfe, wurden bisher selten systematisch einbezogen. Allerdings gibt es an durchaus entscheidenden Stellen bereits einen soliden „Einstieg“ in die regelhafte institutionelle und personelle Einbeziehung der Patientenperspektive [1–3].
- ▲ Eine einseitige Instrumentalisierung der Evidenzbasierten Medizin wird von Patientenorganisationen insbesondere dort erlebt, wo sich durch EbM die Ubernahme von Behandlungskosten infrage stellen lasst.
- ▲ Insbesondere bei seltenen Erkrankungen fehlt hufig eine durch Studien und Meta-Analysen belegte Evidenz. Gleichzeitig werden oft weder der Erfahrungsschatz der organisierten Patientenschaft noch die Evidenz der arztlichen Erfahrungen systematisch genutzt, und wenn, dann erfolgt die Evidenzgenerierung in methodisch zum Teil wenig brauchbarer Durchfuhrung.

Keine Frage, die Patienten wunschen nur diejenige Therapie erleben zu mussen, die mit nachvollziehbarer Evidenz auch Erfolg verspricht. Was dies im Kontext von patienten-

tenrelevanten Endpunkten, Lebensqualität und dem Versichertenstatus, der den verfügbaren Mitteleinsatz definiert, im Einzelfall bedeutet, hängt stets auch von einem vertrauensvollen Arzt-Patienten-Verhältnis ab. Das Spannungsfeld zwischen dem wachsenden Einfluss der Evidenzbasierten Medizin in Kliniken und Praxen und der Ernüchterung aufgrund der angeführten Barrieren bestimmt die heutige Versorgungsrealität aus Patientensicht. Patientenorganisationen setzen sich daher heute zunehmend mit diesen Fragestellungen auseinander. Diese Positionierungen der organisierten Patientenschaft stellen allerdings nicht den Gegenstand dieses Beitrags dar.

Der vorliegende Beitrag fokussiert vielmehr auf die Rolle und Bedeutung der Evidenzbasierten Medizin in der **individuellen** Arzt-Patienten-Interaktion, und hier insbesondere auf die **Evidenzvermittlung**. Unter der Berücksichtigung der Tatsache, dass es nicht „die Patientin“ oder „den Patienten“ gibt, werden auf der Basis unterschiedlicher Informationsbedarfe der Patienten Vorschläge für die gemeinsame Umsetzung der Evidenzbasierten Medizin durch Arzt **und** Patient gemacht sowie praktische Hilfen zur Unterstützung des Arzt-Patienten-Gesprächs bei der Evidenzvermittlung gegeben.

Als praktische Handreichungen wurden unter der Rubrik „Werkzeugkasten für die Arzt-Patienten-Kommunikation“ verschiedene Checklisten, Übersichten und Fragespiegel zusammengestellt, die im Anhang VIII dieses Lehrbuches zu finden sind. Sie stehen darüber hinaus auch auf den Internetseiten des Deutschen Netzwerk Evidenzbasierte Medizin e. V. unter der Adresse <http://www.ebm-netzwerk.de/grundlagen/arbeitsmaterialien> zum Download zur Verfügung.

5.2 Facetten der Evidenz in der Arzt-Patienten-Beziehung

Evidenz in der Arzt-Patienten-Beziehung geht über das ärztliche Wissen um nachgewiesene Wirksamkeit und die Fähigkeit, dieses Wissen zu erlangen, weit hinaus. Damit die Evidenzbasierte Medizin tatsächlich eine Bedeutung in der individuellen Arzt-Patienten-Beziehung erlangt, müssen folgende Bedingungen erfüllt sein:

- ▲ Wissen des Arztes um die aktuelle Datenlage zum vorliegenden gesundheitlichen Problem des Patienten.
- ▲ Patienten sollten – angepasst an ihre Fähigkeiten und Bedürfnisse und unterstützt von ihren behandelnden Ärzten – die Möglichkeit erhalten, mit diesen gemeinsam informierte Entscheidungen treffen zu können [4].
- ▲ Berücksichtigung der Individualität, Wertevorstellungen und Präferenzen des Patienten (s. auch Tab. 5.1) bei der Auswahl infrage kommender Maßnahmen.
- ▲ Vermittlung der Evidenz als Entscheidungsgrundlage an den Patienten („Wie kommt die Evidenz zum Patienten?“).
- ▲ Sicherstellung durch den Arzt, dass die durch ihn vermittelte Evidenz vom Patienten tatsächlich verstanden wurde.

5.2.1 Jeder Patient ist anders

Dass jeder Mensch anders ist, mag eine Binsenweisheit sein. Sie trifft jedoch auch und gerade auf Patienten zu und spielt somit im Interaktionsprozess zwischen Arzt und Patient eine große Rolle. Dies gilt insbesondere dann, wenn es um die Vermittlung von Evidenz geht. Nach den Erfahrungen, die die Autoren im Selbsthilfebereich gewonnen haben, können verschiedene Patiententypen unterschieden werden. Für jeden Typus ist bei der Evidenzvermittlung ein unterschiedliches ärztliches Eingehen erforderlich (s. Tab. 5.1).

Die beobachteten Patiententypen sind:

- ▲ Der Vereinfacher,
- ▲ Der zuhörende Verstehler,
- ▲ Der nachlesende Verstehler,
- ▲ Der Selbstbewerter,
- ▲ Der Erfahrung Einholende,
- ▲ Der Zweitmeinung Einholende,
- ▲ Der Solidarische,
- ▲ Der Zweifler.

Diese Einteilung deckt sich teilweise mit der Einteilung einer Untersuchung des Wellcome-Trust zum Verhältnis der britischen Bevölkerung zur Wissenschaft, die auch für die Rezeption medizinischer Informationen von Relevanz ist [5]. Dort wurde die Einteilung der verschiedenen Typen in „confident believers“, „technophiles“, „supporters“, „concerned“, „not sure“ und „not for me“ vorgenommen.

Tab. 5.1: Patiententypen, ihre Charakteristik und Konsequenzen für den Arzt

Patiententyp	Charakteristik	Konsequenzen für den Arzt
Der Vereinfacher	Ist ein besonders „pflegeleichter“ Patient. Da er davon ausgeht, dass der Arzt sowieso gemäß aktuellem medizinischen Wissenstand berät, empfiehlt und behandelt, ist er an Eigenaktivitäten bezüglich der Informationsbeschaffung nicht interessiert. Vielmehr möchte er seine Beteiligung auf das Befolgen der direkten ärztlichen Vorgaben beschränken.	(Kurze) Erläuterungen für den Patienten geben, auch wenn nicht explizit danach gefragt wurde. Besonders wichtig ist dies bei gesundheitlichen Problemen, für die verschiedene Behandlungsalternativen in Frage kommen. (Kurze) Vergewisserung des Verständnisses und der Zustimmung durch den Patienten. Trotzdem eine ständige Gesprächsbereitschaft signalisieren, falls der Patient zu einem späteren Zeitpunkt doch mehr Informationen wünscht.
Der zuhörende Verstehler	Möchte die Inhalte der mündlichen Beratung, Empfehlungen und Behandlung durch den Arzt nachvollziehen können.	Vergewisserung des Verständnisses und der Zustimmung. Ermutigung, Fragen zu stellen.
Der nachlesende Verstehler	Möchte die Sachverhalte nachlesen und verstehen. Möchte sich Gewissheit verschaffen, dass der Arzt tatsächlich die aktuelle und richtige Information vermittelt hat.	Auf geeignete Möglichkeiten der Informationsbeschaffung verweisen (s. dazu Toolbox „Evidenzquellen für den Arzt“ und „Patienteninformationen“ im Anhang des Buches). Ggf. nennen, auf welcher Informationsquelle(n) die eigenen Kenntnisse beruhen (Von Patienten wird beklagt, dass Ärzte ihre Informationsquellen trotz Nachfrage nicht offen legen und sie daher einen interessengeleiteten „Behandlungsdruck“ in eine bestimmte Richtung befürchten [6]).
Der Selbstbewerter	Möchte die Ergebnisse der Behandlung und ihre Auswirkungen bewusst erfahren und selbst bewerten.	Ermutigung des Patienten, ein Patiententagebuch zu führen und die Inhalte ggf. beim Arztbesuch zu besprechen (s. dazu Toolbox „Dokumentationshilfen für Patienten“ im Anhang des Buches). Ggf. auf geeignete Indikatoren hinweisen, an denen die Behandlungswirkungen verfolgt werden können.

Tab. 5.1: Fortsetzung

Patiententyp	Charakteristik	Konsequenzen für den Arzt
Der Erfahrung Einholende	Möchte sich bei anderen Patienten und Patientenorganisationen erkundigen, welche Erfahrungen mit der Behandlung und/oder dem Behandler vorliegen.	Hinweis auf Patientenorganisationen am Empfang der Praxis (s. dazu Toolbox „Erfahrungsaustausch für Patienten“ im Anhang des Buches). Ggf. schriftliches Einholen des Patienteneinverständnisses zur Kontaktvermittlung mit einem anderen Patienten bzw. Vermittlung eines geeigneten Kontakts (z.B. über eine praxis-interne Kontaktkartei).
Der Zweitmeinung Einholende	Steht lieber „auf zwei Beinen“ und braucht die Sicherheit einer zweiten Expertenmeinung.	Signalisieren, dass die Einholung einer zweiten Meinung keinen Vertrauensbruch darstellt. Ggf. Austausch mit dem Kollegen, bei dem die Zweitmeinung eingeholt wurde.
Der Solidarische	Möchte sein Wissen und die Erfahrungen, die er gemacht hat, auch zum Nutzen anderer Betroffener zur Verfügung stellen.	Hinweis auf Kontaktaufnahme und Mitarbeit in Patientenorganisation oder Selbsthilfegruppe; sofern noch keine diagnosespezifische Gruppe besteht, ggf. auch auf Neugründung (s. dazu Toolbox „Erfahrungsaustausch für Patienten“ im Anhang des Buches). Diese Patienten stellen eine „Goldgrube“ bei der Erschließung und Bündelung von Patientenerfahrungen im Gesundheitswesen dar, die leider noch zu wenig zum Tragen kommt. Patientenerfahrungen sind eine wertvolle Ergänzung des aus externer Evidenz generierten ärztlichen Wissens. Diese Ressource spielt insbesondere bei Seltenen Erkrankungen eine große Rolle und wird u. a. in patientenorientierten Krankheitsbeschreibungen (PKB) erschlossen [7].
Der Zweifler	Hinterfragt ärztliche Erläuterungen und Empfehlungen oft und mehrmals, was vom Arzt als Belastung des Vertrauensverhältnisses empfunden werden kann.	Hier gilt es, wenn möglich, herauszufinden, warum der Patient an den ärztlichen Erläuterungen zweifelt. Oft können zugrunde liegende Ängste durch sachliche Information und/oder persönliches Eingehen gemindert werden. Sofern eine nicht leicht behebbare Störung des Arzt-Patienten-Verhältnisses die Ursache zu sein scheint, kann auch auf die Möglichkeit des Arztwechsels hingewiesen werden.

Bei der in Tabelle 5.1 dargestellten Typologie stellen Mischformen mit einer primären Ausprägung in der Praxis natürlich eher die Regel denn die Ausnahme dar. Welcher primäre Patiententyp vor dem Arzt sitzt, kann

dieser ohne besonderen Aufwand auf folgende Weise herausfinden: Nach dem ersten Besuch kann der Arzt auf vorbereitete Fragenchecklisten, Linklisten, Adressen und Anlaufstellen verweisen, die am Empfang der

Praxis ausliegen. Ob und wie Patienten diese Angebote nutzen, machen beim nächsten Arztbesuch deutlich, welche typbedingten Besonderheiten der Arzt berücksichtigen sollte.

5.2.2 Wie kommt die Evidenz zum Patienten?

Hat der Arzt für eine spezielle gesundheitliche Fragestellung die Evidenzlage geklärt, so steht er vor dem Problem: Wie sage ich es meinem Patienten und zwar so, dass er es auch nachvollziehen und verstehen kann? Ob die Evidenz, die der Arzt seinem Patienten als Entscheidungsgrundlage vermittelt, auch wirklich ankommt, ist von verschiedenen Faktoren abhängig:

- ▲ Kommunikation der tatsächlichen Evidenz,
- ▲ Berücksichtigung patientenrelevanter Endpunkte,
- ▲ Kommunikation der Evidenz in einer für Patienten verständlichen Form,
- ▲ Kommunikationsstil und Empathiefähigkeit des Arztes für die individuelle Patientensituation.

Diese Faktoren sollen nachfolgend erläutert werden:

Kommunikation der tatsächlichen Evidenz

Bei der Darstellung von Behandlungsoptionen ist immer die tatsächliche Evidenzlage zu berücksichtigen. Wenn diese mehrere Schlüsse zulässt, z.B. ob die Methode nachgewiesener Wirksamkeit sowohl kurativ als auch palliativ zum Einsatz kommen kann, muss dies klar dargestellt werden. Das nachfolgende Beispiel aus dem Alltag einer Patientenorganisation soll dies verdeutlichen:

Szenario

Ein Patient hat aufgrund einer lebensbedrohlichen Erkrankung bezüglich der Entscheidung

über die vorgeschlagene Maßnahme einen zweiten Experten konsultiert. Beide Experten empfehlen die gleiche Maßnahme, jedoch stellte ein Arzt die empfohlene Maßnahme als kurativ, der zweite konsultierte Experte diese als palliativ dar. Für den Patienten bedeutete dies entweder „Es gibt noch Hoffnung auf Heilung“ oder „Es ist keine Aussicht auf Heilung; nur noch eine Leidensminderung ist möglich“. Diese Diskrepanz stürzte den Patienten in eine tiefe Krise. Er war der Verzweiflung näher als der evidenzbasierten Behandlungsentscheidung. Und dies nur wegen eines einzigen Wortes, trotz Anwendung der EbM und trotz übereinstimmendem Expertenrat, zumindest was die tatsächlich empfohlene Maßnahme selbst anbelangte.

Ob eine Vermeidung derart abschließend kategorisierender Begriffe wie „palliativ“ oder „kurativ“ erwogen werden sollte und gerechtfertigt ist, hängt nicht zuletzt auch vom Typ des Patienten ab (vgl. Typologie in Tab. 5.1).

Berücksichtigung patientenrelevanter Endpunkte

Gelegentlich werden bei einer infrage kommenden Therapie ausschließlich positive Entwicklungen von Surrogatparametern (z.B. Senkung des Blutfettspiegels) als Nutzen für die infrage kommenden Behandlungsmöglichkeiten angeführt, weil sie in ihrem zeitlichen Verlauf einfacher und schneller zu beobachten sind als die tatsächlichen patientenrelevanten Endpunkte. Solche Surrogatparameter reichen jedoch als Nutzenangabe nicht aus. Ärzte sollten die Darstellung des erwarteten Nutzens auf Parameter beziehen, die für den Patienten bedeutsam sind [8–11], also auf:

- ▲ höhere Lebenserwartung,
- ▲ weniger Beschwerden und Komplikationen,
- ▲ höhere gesundheitsbezogene Lebensqualität,
- ▲ Erleichterung der sozialen Teilhabe,
- ▲ weniger Aufwand,

- ▲ weniger Inanspruchnahme von Leistungen des Gesundheitssystems,
- ▲ Erleichterung des Selbstmanagements der Erkrankung z.B. bei der Handhabung und Applikation von Medikamenten.

Kommunikation der Evidenz in einer für Patienten verständlichen Form

Bei der Kommunikation der Evidenz kommt es auf die Vermittlung des Ausmaßes des Nutzens und/oder möglichen Schadens bzw. der Risiken einer Maßnahme an, die gegeneinander abgewogen werden müssen.

Die wissenschaftliche Datenlage stellt Wahrscheinlichkeiten für den Nutzen bzw. die Risiken einer Maßnahme dar. Patienten haben aber aufgrund ihrer individuellen Betroffenheit verständlicherweise ein Bedürfnis nach Sicherheit. Sie erwarten eine „Ja/Nein“-Entscheidung [12]. Dies stellt in Bezug auf die Kommunikation von Evidenz besonders hohe Anforderungen an den Arzt.

Die unterschiedliche Darstellung identischer Sachverhalte (Framing) hat ebenfalls einen Einfluss auf die Entscheidung. So führt z.B. eine positive Darstellung des Risikos von Therapien (z.B. „99% der Patienten haben keine Komplikationen.“) eher dazu, dass sich Patienten für diese Behandlung entscheiden, als eine negative Darstellung (z.B. „Komplikationen treten bei einem von 100 Patienten auf.“) [13]. Ärzte sollten ihren Patienten immer beide Darstellungen kommunizieren, um eine Entscheidungsverzerrung durch dieses Framing möglichst auszuschließen. Dies sollte stets in Verbindung mit einer verständlichen Begründung für die Präferenz des Arztes erfolgen.

Numerische Angaben in medizinischen Sachverhalten zu verstehen und vor dem Kontext der individuellen Fragestellung richtig zu interpretieren, stellt für viele Patienten eine Überforderung dar. Der Arzt sollte daher in einer möglichst verständlichen Form, z.B. unter Angabe von absoluten Zahlen mit nachvollziehbaren Bezugsgrößen, eine „Über-

setzung“ numerischer Evidenzangaben für seine Patienten vornehmen. Als Beispiel soll das oft zitierte Entscheidungsbeispiel zur Teilnahme am Brustkrebs-Screening dienen: Die Anzahl der durch Brustkrebs bedingten Todesfälle bei jeweils 1000 Frauen beträgt mit Mammographie-Screening 3 und ohne Mammographie-Screening 4. Wenn man die Relative Risiko-Reduktion ausrechnet, die 25% beträgt, könnte das zur missverständlichen Interpretation führen, dass etwa von 100 Frauen, die am Mammographie-Screening teilnehmen, 25 gerettet werden können. Tatsächlich kann aber durch Teilnahme am Mammographie-Screening ein Todesfall mehr verhindert werden gegenüber den Frauen, die nicht teilnehmen [14–16].

Bei der Kommunikation der Evidenz sollte auch darauf eingegangen werden, ob beim betreffenden Krankheitsbild in Bezug auf die Behandlung Unterschiede des Geschlechts, der ethnischen Herkunft oder sonstige Faktoren eine Rolle spielen können bzw. bekannt sind.

Kommunikationsstil und Empathiefähigkeit des Arztes für die individuelle Patientensituation

Ob Sachinformationen vom Patienten aufgenommen und verstanden werden, ist neben ihrer richtigen und vor allem verständlichen Darstellung auch vom Kommunikationsstil des Arztes und seinen empathischen Fähigkeiten abhängig. Balint [17] fordert als Grundlage einer geglückten Arzt-Patienten-Kommunikation, dass Ärzte lernen müssen (und können!), die Beziehungen zu ihren Patienten heilsam-menschlich zu gestalten. Dazu gehört nicht nur die kommunikative Empathie des Arztes, sondern auch ein störungsfreies und beruhigendes Umfeld – also möglichst keine Aufklärungs- und Informationsgespräche auf Klinikfluren (die es leider immer noch gibt).

Eine schwere oder gar lebensbedrohliche Erkrankung trifft nie die Patienten allein,

sondern auch den familiären und gesellschaftlichen Kontext, in dem sie sich bewegen. Das löst starke Ängste aus, die so prägend sein können, dass Sachinformationen, wie die Wahrscheinlichkeit des erwarteten Nutzens einer vorgeschlagenen Maßnahme, weder aufgenommen noch verstanden werden können. Ehe solche Sachinformationen gegeben werden, muss sich der Arzt mit diesen Ängsten seines Patienten verständnisvoll auseinandersetzen und so eine Vertrauensbeziehung herstellen. Erst dann können Sachinformationen diskutiert und auf ihrer Basis gemeinsam Therapieentscheidungen getroffen werden.

In Fällen, wo z.B. die gesicherte Evidenz darin besteht, nichts zu tun, kann dies dem Patienten unter Umständen schwer vermittelbar sein. Gerade bei schweren Erkrankungen brauchen Patienten das Gefühl, dass alles getan wird, um ihre gesundheitliche Situation zu verbessern. Hier gilt es, einen geeigneten Weg zu finden, um den Patienten zu beruhigen und ihm soweit wie möglich sachlich begründet und einfühlsam seine Ängste zu nehmen [18].

5.3 Patientenerfahrungen – auch eine Evidenzquelle für die Medizin

Die Wahrscheinlichkeit, Belege nachgewiesener Wirksamkeit hoher Evidenzklassen im Falle häufiger Erkrankungen oder Erkrankungen von gesellschaftlicher Relevanz zu finden, ist ungleich höher als bei seltenen Erkrankungen. Hier stellen strukturierte Erfahrungsberichte von Patienten derzeit nicht selten die einzige verfügbare Evidenzquelle für versorgungsrelevante Fragestellungen dar.

Der Erschließung und Strukturierung von Patientenerfahrungen wird vonseiten der Patientenorganisationen eine große Bedeutung beigemessen. Ansätze zur systemati-

schen Nutzung dieser Ressourcen existieren bereits [19]. Verschiedene Gruppen und Organisationen wie die Allianz Chronischer Seltener Erkrankungen ACHSE (<http://www.achse-online.de>) und die Arbeitsgruppe Methodik im Patientenforum (<http://www.versorgungsleitlinien.de>) beschäftigen sich mit der Systematisierung und der Schaffung methodischer Grundlagen zur Erschließung und Strukturierung des Erfahrungswissens der Patientenorganisationen [20]. Hierbei ist es für Patientenorganisationen besonders wichtig, auch Patiententypen einzubeziehen, die sich ansonsten nicht mit Fragen der Informationsgewinnung und Arzt-Patienten-Beziehung auseinandersetzen, also außer dem „Solidarischen“ auch alle weiteren „Patiententypen“ anzusprechen. Dies erst ermöglicht eine umfassende Sicht auf die tatsächliche Versorgungssituation und minimiert den „Auswahl-Bias“ bei der Erhebung der Patientenerfahrungen.

In diesem Zusammenhang sei darauf hingewiesen, dass im angelsächsischen Raum unter der Internetadresse <http://www.dipex.org> eine äußerst umfangreiche Datenbank von Patientenerfahrungen bei unterschiedlichen Erkrankungen und Therapieformen existiert.

5.4 Zusammenfassung

Die Frage „Welche Evidenz will der Patient?“ kann nicht pauschal beantwortet werden. Sie ist immer vom individuellen Patienten, also seiner speziellen Erkrankung, seinem Umfeld, seinen Sichtweisen und Bedürfnissen abhängig. Dies trifft auch für die Vermittlung von Evidenz zu. Bei der Beantwortung spezieller gesundheitlicher Fragestellungen im individuellen Fall muss der Arzt seine Analyse der aktuellen Datenlage auch mit den Erwartungen, Wertvorstellungen und Präferenzen des Patienten abgleichen. Dies ist nur in einer von Vertrauen geprägten

Atmosphäre der Arzt-Patienten-Beziehung möglich. Die Übernahme von mehr Eigenverantwortung setzt auch voraus, dass Patienten die erforderlichen Informationen erhalten, um auf deren Basis mit ihren Ärzten gemeinsam verantwortete Entscheidungen treffen zu können. Hierzu sollten Patienten zumindest eine ungefähre Vorstellung davon haben, wie Evidenz generiert und im individuellen Fall interpretiert wird. Dieses Informationsbedürfnis wird von Ärzten oft unterschätzt. Es gibt in Deutschland bereits Ansätze, die Grundsätze und Methodiken der Evidenzbasierten Medizin auch an medizinische Laien (z.B. durch spezielle Trainingsmodule in wissenschaftlicher Kompetenz [21]) und an in der Patientenberatung tätige Personen zu vermitteln (z.B. Curriculum Patientenberatung [22]). Jedoch kann natürlich nicht jeder interessierte Patient an einem solchen Training teilnehmen. Vor diesem Hintergrund ergeben sich heute neue Erwartungen und Anforderungen an die ärztliche Kompetenz [23], die insbesondere die Bereiche der medizinischen, wissenschaftlich-methodischen, psychosozialen und kommunikativen Fähigkeiten beinhalten.

Die Kooperation mit Patientenorganisationen, für die das Thema Evidenzbasierte Medizin ebenfalls von wachsendem Interesse ist, kann in diesem Prozess eine unterstützende Wirkung haben. Schon jetzt gibt es bei verschiedenen Projekten, so z.B. bei der Übertragung evidenzbasierter Nationaler VersorgungsLeitlinien in PatientenLeitlinien [2], eine enge Zusammenarbeit mit Patientenorganisationen, vertreten durch das Patientenforum bei der Bundesärztekammer [24]. Was Patienten vor dem Hintergrund der praktischen Umsetzung der Evidenzbasierten Medizin wirklich wollen, können nur sie selbst definieren und individuell sowie über ihre Organisationen einfordern. In Patientenorganisationen wird derzeit an den Voraussetzungen gearbeitet, diese Bedürfnisse in einer systematischen und strukturierten

Form zu erfassen und in den Dialog mit medizinischen Berufsgruppen und anderen Adressaten einzubringen.

5.5 Werkzeugkasten für die Arzt-Patienten-Kommunikation

Eine Frage, die an Patientenorganisationen, den Patienteninformationsdienst der Ärzteschaft (www.patienten-information.de) und an das Deutsche Netzwerk Evidenzbasierte Medizin vonseiten der Patienten immer wieder heran getragen wird, ist: „Wo finde ich einen Arzt, der nach den Prinzipien der Evidenzbasierten Medizin arbeitet?“ Da es keine speziell registrierten „EbM-Ärzte“ gibt, kann diese Frage nicht beantwortet werden. Eine Hilfe für individuelle Patienten und Patientenorganisationen ist jedoch die Beantwortung der Frage „Woran **erkenne** ich, ob ein Arzt nach den Prinzipien der Evidenzbasierten Medizin arbeitet?“. Erste Schritte in diese Richtung wurden bereits gemacht, etwa durch die von Patientenvertretern und dem Ärztlichen Zentrum für Qualität in der Medizin entwickelte Checkliste „Woran erkennt man eine gute Arztpraxis?“ [25] oder die Artikelserie „Medizinische Informationen bekommen und richtig verstehen“ von Klemperer und Lang [26]. Dieser Weg sollte konsequent weiter verfolgt werden.

Jeder Arzt – egal ob in Klinik und Praxis tätig – kann seinen Patienten entsprechende Materialien zur Verfügung stellen, die die Aufgabe haben,

- ▲ das Arzt-Patienten-Gespräch zu unterstützen (✳ Toolbox „Unterstützung des Arzt-Patienten-Gesprächs“ in Anhang VIII des Buches),
- ▲ dazu ermutigen, den Erfahrungsaustausch mit gleich Betroffenen zu suchen (✳ Toolbox „Erfahrungsaustausch mit gleich Betroffenen“ in Anhang VIII des Buches),

- ▲ Patienten über verlässliche zusätzliche Informationsangebote zu informieren (✳ Toolbox „Verlässliche Informationsangebote im Internet“ in Anhang VIII des Buches),
- ▲ Patienten eine Hilfestellung bei der Einschätzung der Qualität von Informationen, insbesondere im Internet, zu geben (✳ Toolbox „Hilfestellung bei der Einschätzung der Qualität von Informationen im Internet“) in Anhang VIII des Buches.

Literatur

- [1] Gemeinsamer Bundesausschuss (G-BA) (2005), Patientenbeteiligung im Gemeinsamen Bundesausschuss. http://www.g-ba.de/cms/front_content.php?idcatart=177&lang=1&client=1 (11.07.06)
- [2] Sänger S, Legemann M, Kopp I, Patientenbeteiligung an der NVL-Asthma – ein Werkstattbericht. *Z Arztl Fortbild Qualitätssich.* (2006), 100(6), 447–452
- [3] Weingart O et al., Patientenbeteiligung in der Cochrane Collaboration-Barrieren, Erfahrungen und Konzepte aus deutscher Sicht. *Z Arztl Fortbild Qualitätssich.* (2005), 99(6), 367–371
- [4] Härter M, Partizipative Entscheidungsfindung (Shared Decision Making) – ein von Patienten, Ärzten und der Gesundheitspolitik geforderter Ansatz setzt sich durch. *Z Arztl Fortbild Qualitätssich.* (2004), 98(2), 89–92
- [5] Office of Science and Technology, Wellcome Trust (2000), Science and the public. A review of science communication and public attitudes to science in Britain. <http://www.wellcome.ac.uk/assets/wtd003419.pdf> (11.07.06)
- [6] Stötzner K, Einbindung von Patienten und ihren Anliegen in die Evidenzbasierte Medizin. *Z Arztl Fortbild Qualitätssich.* (2001), 95, 529–531
- [7] Allianz Chronischer Seltener Erkrankungen. Wissen verbreiten: Patientenorientierte Krankheitsbeschreibungen. 2006. http://www.achse-online.de/projekte/projekte_04.php (11.06.06)
- [8] Fischer G, Die Sicht von Arzt und Patient. *Dtsch Med Wochenschr.* (2006), 131, 8–11
- [9] Sawicki P, Patientenrelevante Endpunkte – Stand der Diskussion im Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). *Dtsch Med Wochenschr.* (2006), 131, 16–20
- [10] Sänger S et al. (2006), Manual Patienteninformation. Empfehlungen zur Erstellung evidenzbasierter Patienteninformationen. BoD, Norderstedt. http://www.patienten-information.de/content/download/manual_patienteninformation_04_06.pdf (23.05.06)
- [11] Windeler J, Nutzen und Nutzenbewertung. *Dtsch Med Wochenschr.* (2006), 131, 12–15
- [12] Powell D, Leiss W (1997) Mad cows and mother's milk: case studies in risk communication. McGill-Queen's Univ. Pr., Montreal
- [13] Klemperer D, Rosenwirth M, Bertelsmann Stiftung, Universität Bremen, Zentrum für Sozialpolitik (2005), Shared Decision Making: Konzept, Voraussetzungen und politische Implikationen. Chartbook. 2. Auflage. <http://www.patient-im-mittelpunkt.de/2005-07chartbookSDM.pdf> (11.07.06)
- [14] Gigerenzer G (2004) Das Einmaleins der Skepsis. Über den richtigen Umgang mit Zahlen und Risiken. Berliner Taschenbuch-Verlag, Berlin
- [15] Norman CD, Skinner HA, eHealth literacy: Essential skills for consumer health in a networked world. *J Med Internet Res.* (2006), 8(2), e9
- [16] Steckelberg A et al., Kriterien für evidenzbasierte Patienteninformationen. *Z Arztl Fortbild Qualitätssich.* (2005), 99(6), 343–351
- [17] Balint M (2001) Der Arzt, sein Patient und die Krankheit, 10. Aufl. Klett-Cotta, Stuttgart
- [18] Lang B, Sänger S, Evidenzbasierte Medizin für den Bürger. Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz. (2005), 48(6), 679–684
- [19] Emmerich M, Zusammenfassung einer ZSD-Selbstbefragung. *Retina akt.* (2005), 2, 29–30
- [20] Brunsmann F et al., Patientenbeteiligung – Diagnosespezifische Erfahrungen einbringen. *Dt Arztebl.* (2004), 101(28-29), A2026–9

- [21] Training in wissenschaftlicher Kompetenz für Patienten- und Verbrauchervertreter (2006). http://gesundheit.chemie.uni-hamburg.de/cgi-bin/newsite/index.php?page=page_59 (11.07.06)
- [22] Dierks ML et al. (2004) Curriculum Patientenberatung. Informationen recherchieren, bewerten, vermitteln, 2. Aufl. Ärztliches Zentrum für Qualität in der Medizin, Berlin 2004. http://www.patienten-information.de/content/projekte/kontinuierlich/download/curriculum_ebm_patinfo_akt.pdf (12.12.06)
- [23] Ollenschläger G, Die Kunst, heutzutage ein guter Arzt zu sein. *Ärztebl Bad-Würt.* (2004), 59, 111–114
- [24] Sänger S et al., Mehr Power für Patienten: Patientenforum – Gemeinsame Plattform von Ärzteschaft und Patientenselbsthilfe. *Z Arztl Fortbild Qualitätssich.* (2001), 95, 531–533
- [25] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ) (2005), Woran erkennt man eine gute Arztpraxis? Checkliste für Patientinnen und Patienten. http://www.patienten-information.de/content/informationsqualitaet/checkliste_arztbesuch/view (08.12.05)
- [26] Klemperer D, Lang B (2006), Der große Ärzte-Check. Machen Sie sich schlau. Medizinische Informationen bekommen und richtig verstehen. STERN-Serie. http://kurse.fh-regensburg.de/kurs_20/kursdateien/alle/stern.html

6 Welche Evidenz braucht das System?

Reinhard Busse, Bernhard Gibis

Da die Evidenzbasierte Medizin vornehmlich auf die Entscheidungsfindung in der täglichen, klinischen Medizin abzielt, stellt sich die Frage, ob und in welcher Form ihre Grundkomponenten auf die Versorgung der gesamten Bevölkerung bzw. das Gesundheitssystem übertragen werden können. International wird dies unter dem Schlagwort „evidenzbasierte Gesundheitsversorgung“ (englisch: Evidence-based health care) diskutiert.

Entscheidungen auf Systemebene sind vielfältig: Sollen eine bestimmte diagnostische oder therapeutische Technologie oder ein bestimmtes Arzneimittel in den Leistungskatalog eingeführt werden? Welche vorgeblich qualitätssichernden Maßnahmen verbessern tatsächlich die Gesundheitsversorgung? Sind im Ausland eingeführte Screening-Programme für die Anwendung im hiesigen Gesundheitssystem geeignet? etc. – dies sind allesamt Fragen, für die tagtäglich eine informierte Entscheidung durch dazu berufene Entscheidungsträger zu treffen ist.

Die klassische Trias der EbM aus Wunsch des Patienten, Erfahrung des (zumeist ärztlichen) Leistungserbringers und externer Evidenz kann auf die Anforderungen auf Systemebene nicht unmittelbar übertragen werden und Anwendung finden. Was sind die Wünsche der Patienten in diesem Zusammenhang? Was entspricht der persönlichen Erfahrung des Arztes bei Entscheidungen auf Systemebene? Wie die Evidenzbasierte Medizin insgesamt, hat die evidenzbasierte Gesundheitsversorgung die am weitesten reichenden Fortschritte im angelsächsischen Sprachraum gemacht. Dort wird ihr eine Trias aus Erkenntnissen aus hochwertigen wis-

senschaftlichen Untersuchungen, zur Verfügung stehenden Ressourcen und „Werten“ zugeordnet.

Am Beispiel des Mammographie-Screenings soll verdeutlicht werden, welcherart diese Informationen sein können und wie letztlich trotz aller Information die Einführung eines umfassenden Screening-Programms eine „Werteentscheidung“ bleibt. Ebenfalls zur Darstellung kommen soll, wie die Entscheidungsfindung in diesem Fall in Deutschland erfolgte. An diesem Beispiel werden die Chancen und Limitationen der evidenzbasierten Gesundheitsversorgung in der Planung der Gesundheitsversorgung aufgezeigt.

6.1 Methodische Überlegungen bezüglich evidenzbasierter Systementscheidungen

Zunächst aber folgen methodische Überlegungen (insbesondere zur ersten Säule der Trias), die für die Nutzung von Daten aus klinischer Forschung für bevölkerungsmedizinische bzw. das Gesundheitssystem betreffende Entscheidungen wichtig, aber oftmals nicht bekannt sind – bzw. vergessen werden.

In der klinischen Forschung gelten randomisierte kontrollierte Studien (RCTs) als Goldstandard der Beweisführung, da Störgrößen so weit wie möglich ausgeschaltet bzw. angeglichen werden. Zu diesen Störgrößen gehört bei der klinischen Forschung auch das Gesundheitssystem bzw. der Behandlungsstil verschiedener Leistungserbringer. Deswegen legen detaillierte Studien-

protokolle fest, wieviele Leistungen von wem zu erbringen sind, wobei dies zumeist in Universitätskliniken unter bestmöglichen Bedingungen erfolgt. Die unter solchen „idealen“ Umständen gemessene Wirksamkeit eines Arzneimittels, einer Therapieform oder anderer Technologien wird als „Efficacy“ (auf Deutsch etwa „Effektivität unter optimalen Bedingungen“) bezeichnet. Für eine systembezogene Entscheidungsfindung ist aber die Kenntnis um – und Berücksichtigung von – solchen in der klinischen Forschung als Störgrößen betrachteten Einflussfaktoren notwendig, um aufgrund der (erwarteten bzw. zu erreichenden) „Effektivität unter Alltagsbedingungen“ evidenzbasiert entscheiden zu können.

Bleiben wir aber noch einen Moment bei den RCTs: Um die interne Validität der Ergebnisse sicherzustellen, schließen RCTs große Gruppen von Patienten aus, die unter Alltagsbedingungen die Therapie bekommen würden. So wird fast immer eine untere und eine obere Altersgrenze für Studienteilnehmer definiert und Frauen (insbesondere im gebärfähigen Alter) sowie Patienten mit Komorbidität werden häufig ganz ausgeschlossen. Auch die exakte Festlegung klinischer Parameter als Einschluss- bzw. Ausschlusskriterien führt zum Ausschluss von relevanten Patientengruppen. So stellten Hlatky und Kollegen (1988) [1] fest, dass von ihren 719 Patienten, die sich einer Bypass-Operation unterzogen hatten, nur 4%, 9% bzw. 12% die Einschlusskriterien der in renommierten Journals publizierten Studien (Veterans Administration Cooperative Study, European Cooperative Surgery Study, Coronary Artery Surgery Study) erfüllt hätten.

Ein zusätzlicher Faktor, der bei RCTs zu beachten ist, ist der oft nicht begründete Ausschluss bzw. die Nicht-Teilnahme von potenziellen Studienteilnehmern, die die Einschlusskriterien erfüllen. Remington (1989) [2] hat darauf hingewiesen, dass bei der großen Hypertonie-Studie des britischen

Medical Research Council mehr als die Hälfte der 46000 geeigneten Patienten ausgeschlossen wurden bzw. nicht teilnahmen, wobei dies anscheinend diejenigen mit im Schnitt schlechterer Prognose waren, da die Mortalität auch der Kontrollgruppe deutlich niedriger als erwartet ausfiel.

Kontrovers wird die Frage diskutiert, ob „echte“ experimentelle Studien – d.h. Studien, bei denen der Forscher sowohl die Intervention als auch die Zuordnung der Studienteilnehmer prospektiv kontrollieren kann – gegenüber nicht-experimentellen systematisch zu einer Unter- bzw. Überschätzung des tatsächlichen Ergebnisses führen. Lange Zeit galt als gegeben, dass nicht-randomisierte Studien zu einer konsistenten Überschätzung von Behandlungseffekten führen [3–8]; eine genauere Analyse der Daten zeigt aber, dass dies vor allem auf Studien mit historischen Kontrollgruppen zutrifft. Außerdem weisen innerhalb der Gruppe der RCTs die nicht-verblindeten Studien deutlich höhere Effekte auf.

Ein systematischer Review von 18 Publikationen mit Daten aus randomisierten und nicht-randomisierten Studien fand keine systematischen Effektunterschiede zwischen RCTs und „Beobachtungsstudien“, sofern bei letzteren die gleichen Aus- und Einschlussbedingungen wie bei RCTs angewendet wurden [9].

Insgesamt lässt sich festhalten, dass vor allem folgende Faktoren die Effektgröße in RCTs erhöhen:

- ▲ Auswahl der teilnehmenden Leistungsanbieter, die meist auf ihrem Gebiet spezialisiert sind,
- ▲ Studienbedingungen und
- ▲ ggf. Ausschluss von älteren Patienten und Patienten mit Komorbidität.

Dagegen erhöhen folgende Faktoren die Effektgröße in Beobachtungsstudien:

- ▲ fehlender Placeboeffekt,
- ▲ Nicht-Verblindung (d.h. ggf. Höherbewertung des Effektes in der Interventionsgruppe),

- ▲ Möglichkeit zur Zuteilung zu einer für den individuellen Patienten angemesseneren Therapie,
- ▲ Berücksichtigung von Patientenpräferenzen und
- ▲ bei historischen Kontrollen andere in zwischen hinzugetretene Fortschritte.

Zum Zwecke evidenzbasierter Gesundheitsversorgung ist es oftmals angemessen, die Daten kontrollierter Studien mit Routinedaten zu ergänzen [10], um z.B. die Repräsentativität der Studienpopulationen mit der tatsächlichen Bevölkerung hinsichtlich demographischer Parameter oder ihres Nutzungsverhaltens der diversen Versorgungsangebote festzustellen (nicht hingegen, um den Wirksamkeitsnachweis zu ersetzen!). Routinedaten werden auch „Sekundärdaten“ genannt, da sie primär zu anderen Zwecken erhoben wurden. Dies sind in erster Linie

(Leistungs-)Daten z.B. von Krankenkassen, aber auch Daten aus klinischen Studien, Bevölkerungssurveys oder etwa von der BQS (Bundesgeschäftsstelle Qualitätssicherung), die zu diesem Zweck neu ausgewertet werden.

Dass die aus klinischen Studien stammenden Daten z.T. sehr stark von bevölkerungsbezogenen Daten abweichen, sei in Abb. 6.1 beispielhaft für Komplikationen nach radikaler Prostatektomie gezeigt. Die Daten zur Effektivität entstammen einer Analyse von Medicare-Routinedaten, die zur Efficacy einer Meta-Analyse mit Daten aus RCTs.

Neben verschiedenen prädisponierenden Faktoren auf der Patientenseite (Alter, Komorbidität etc.) ist die – aufgrund Erfahrung, technischer Ausstattung oder Motivation – unterschiedliche Qualität der Leistungserbringer bei der Durchführung der Maßnah-

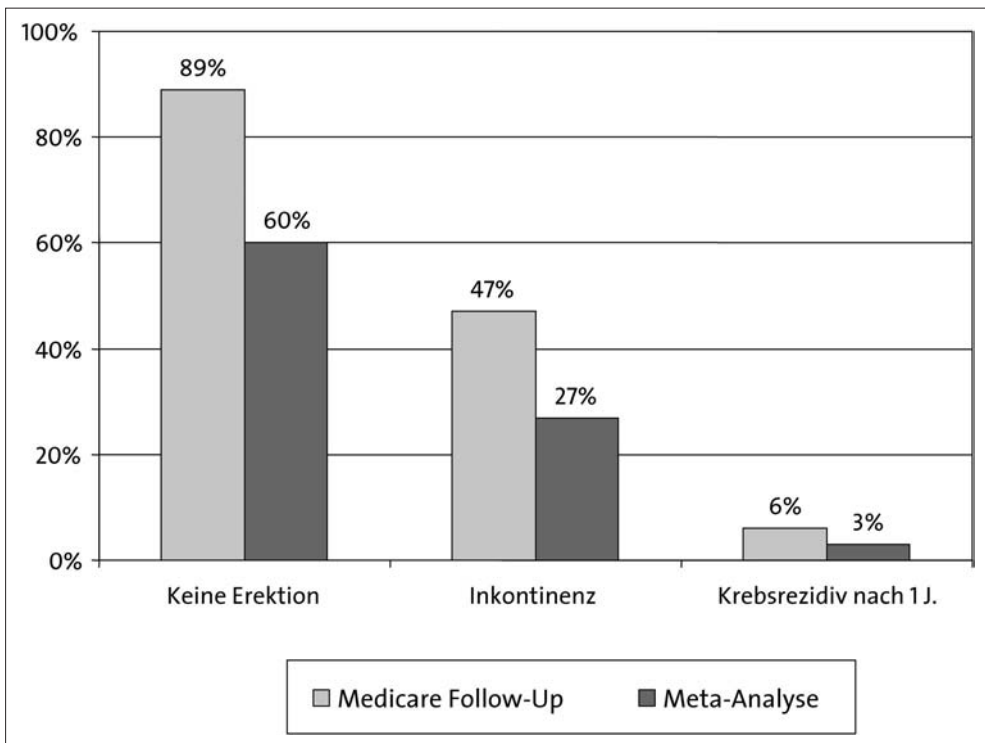


Abb. 6.1: Effektivität unter Alltagsbedingungen versus „Efficacy“ – Das Beispiel radikale Prostatektomie (eigene Darstellung nach [11])

me ein wesentlicher Faktor für den Unterschied zwischen Efficacy und Effektivität. Aus Sicht des Gesundheitssystems ist „quality of care that component of the difference between efficacy and effectiveness that can be attributed to care providers, taking account of the environment in which they work“ [12].

6.2 Bevölkerungs- bzw. systembezogene Effektivität: viele Faktoren und ein Modell

Die für Efficacy oder „effectiveness“ gängigen Definitionen gehen (explizit oder implizit) davon aus, dass der Unterschied nur für die Personen relevant ist, die sich der entsprechenden Maßnahme tatsächlich unterziehen – und auch die Daten in Abb. 6.1 beziehen sich ja nur auf Patienten. Tugwell und Kollegen (1984) [13] waren die ersten, die den Versuch unternommen haben, die Konzepte von Efficacy und Effektivität auf die Bevölkerung zu übertragen. Sie führten den Begriff der „community effectiveness“ ein, den sie wie folgt operationalisierten:

$$\text{Community effectiveness} = \text{efficacy} \times \text{diagnostic accuracy} \times \text{health professional compliance} \times \text{patient compliance} \times \text{coverage}$$

Die „Formel“ führt einige Variablen auf, die zum Unterschied von Efficacy und Effektivität beitragen. Die Liste ist jedoch unvollständig und der mathematische Zusammenhang in der vereinfachten Form zweifelhaft. Analysiert man die genannten Variablen, gibt es mit „Coverage“ (also Personen mit entsprechendem Versicherungsschutz für die Leistung) einen systemseitigen Faktor, mit „Compliance“ einen patientenseitigen Faktor und mit diagnostischer Genauigkeit und Compliance (mit entsprechenden Leitlinien; „Do the right thing“) zwei arztseitige Fakto-

ren. Damit fehlen als wesentliche Faktoren zumindest der patientenseitige Zugang zum Leistungserbringer bzw. zur Leistung und von den arztseitigen Faktoren die (technische) Qualität der Leistungserbringung („Do the thing right“ [14]).

Außerdem unterstellt ihre Formel einen festen, für alle Patienten gleichen Wert von Efficacy. Bei den meisten medizinischen Leistungen hängt der zu erwartende gesundheitliche Effekt einer Maßnahme jedoch stark vom individuellen Fall ab. Dies ist das Tätigkeitsfeld der Angemessenheits-Forschung („Appropriateness Research“). Als angemessen gilt dabei, wenn der „expected health benefit exceeds the expected negative consequences by a large enough margin to justify performing the procedure rather than other alternatives“ [15]. Das zur Untersuchung der Angemessenheit häufig angewendete so genannte Rand-Verfahren beruht in seinem Kern auf durch die Delphi-Methode zustande gekommenem Experten-Konsens. Diese bewerten eine große Anzahl möglicher Indikationen (= Fallbeschreibungen mit klinischen Parametern einschließlich Komorbidität sowie ggf. Patientenalter und -geschlecht) daraufhin, ob bei ihnen die Maßnahme als angemessen, unangemessen oder als „unsicher“ gelten kann. In einer weiteren Stufe werden diese Bewertungen dann auf tatsächliche Patientenkollektive angewendet, sodass ermittelt werden kann, bei wievielen Patienten eine Maßnahme angemessen bzw. unangemessen war.

Seit Beginn dieser Forschungsrichtung wurde auch untersucht, ob erhöhte Raten durchgeführter Prozeduren mit erhöhten Raten unangemessenen Einsatzes zusammenhängen. Dass die ermittelten Angemessenheitsraten bei hohen und niedrigen Durchführungsraten entgegen der Erwartung recht ähnlich sind, d.h. der Zusammenhang allenfalls schwach negativ ist, hat Wennberg (1987) [16] „The paradox of appropriate care“ genannt.

Aufgrund der Schwäche existierender Ansätze, die Efficacy einer Maßnahme auf der Bevölkerungsebene – also den durch eine medizinische Technologie maximal zu erzielenden bevölkerungsgesundheitlichen Nutzen – abschätzen zu können, soll im Folgenden ein neues Konzept kurz skizziert werden. Dieses integriert die Überlegung der Angemessenheit und ermöglicht durch die Berücksichtigung verschiedener Variablen des Systems, der Leistungserbringer und der Patienten eine Abschätzung der Bevölkerungseffektivität, also des unter gegebenen Umständen tatsächlich zu erzielenden bevölkerungsgesundheitlichen Nutzens.

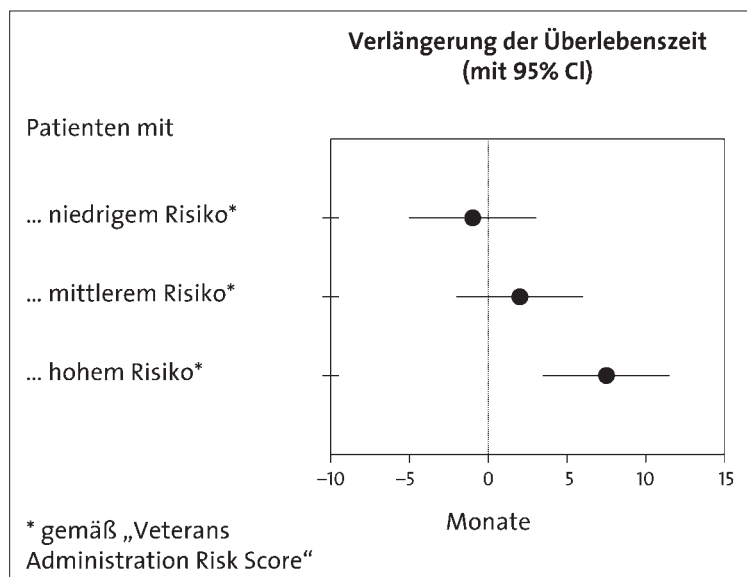
Diese Konzeption dürfte für evidenzbasierte Versorgungsentscheidungen relevant sein, da dabei die bevölkerungs- bzw. systemweiten Auswirkungen einer (neuen) Technologie im Mittelpunkt stehen. Wichtige diesbezügliche Fragen betreffen die Übertragbarkeit von unter Studienbedingungen gewonnenen Ergebnissen und die Übertragbarkeit von in einem Land gewonnenen Ergebnissen in ein anderes.

Die Bevölkerungs-Efficacy hängt von zwei Faktoren ab: von der Efficacy der Maß-

nahme für den einzelnen Patienten und von der Anzahl von Personen in der Bevölkerung, für die eine vergleichbare Efficacy erwartet werden kann. Den Ausgangspunkt des Konzeptes bilden Ergebnisse randomisierter klinischer Studien (RCTs), welche eine Differenzierung nach Risikoprofil, Schweregrad, Indikation etc. innerhalb des teilnehmenden Patientenkollektives ermöglichen. Leider stellt – neben dem Ausschluss großer Patientengruppen – die Differenzierung in für statistische Zwecke genügend große, für die klinische Ergebnisinterpretation aber relevante Gruppen immer noch eine Schwachstelle der meisten RCTs dar. So stellt sich für ihre Ergebnisse nicht nur die Frage der Übertragbarkeit auf ausgeschlossene Gruppen, sondern auch auf Patienten, die die Einschlusskriterien erfüllen. Am Beispiel der Bypass-Chirurgie (im Vergleich zur medikamentösen Behandlung) stellt Abb. 6.2 das Ergebnis einer Meta-Analyse von Yusuf und Kollegen (1994) [17] in grafisch aufbereiteter Form dar.

Wird das Konzept der Angemessenheit auf diese Daten übertragen, ergibt sich eine „angemessene“ Indikationsstellung nur für

Abb. 6.2: Lebensverlängerung innerhalb von 10 Jahren für Teilnehmer an sieben Studien zum Effekt der Bypass-Chirurgie im Vergleich mit medikamentöser Behandlung, stratifiziert nach niedrigem/mittlerem/hohem Risikoprofil (aus [18] mit Daten von [17])



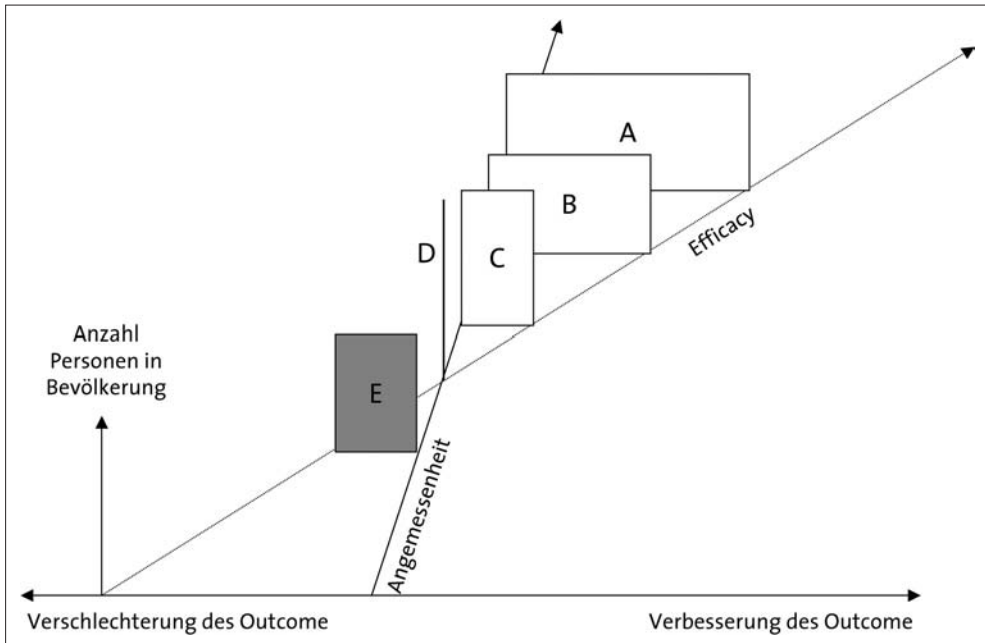


Abb. 6.3: Modell zur Abschätzung der Bevölkerungs-„Efficacy“ einer Maßnahme unter Annahme von fünf differenzierbaren Risiko- bzw. Indikationsgruppen, von denen drei (A, B, C) ein positives Outcome, eine (D) keinen Effekt und eine (E) ein negatives Outcome erwarten können; die Bevölkerungs-„Efficacy“ ergibt sich aus der Summe der weißen Rechtecke abzüglich des grauen Rechtecks

Patienten mit mittlerem oder hohem Risiko, während sie für Patienten mit niedrigem Risiko „unangemessen“ ist, da hier eine Verkürzung der Überlebenszeit zu erwarten ist. Würden hingegen nur die Gesamtergebnisse der Meta-Analyse zugrunde gelegt, ergäbe sich bei positivem Effekt für das Gesamtkollektiv auch für Patienten mit niedrigem Risiko (die ja in die RCTs eingeschlossen waren) fälschlicherweise eine „angemessene“ Indikation.

Solcherlei Studiendaten bilden die Grundfläche in der Konzeption zur Bevölkerungs-Efficacy, während die in Abb. 6.3 dargestellten Höhen über den (beispielhaft) fünf Risiko- bzw. Indikationsgruppen die jeweilige Anzahl von Personen in der Bevölkerung widerspiegeln. Diese Daten müssten durch epidemiologische Untersuchungen in der Bevölkerung gewonnen werden, wenn sie nicht anderweitig abschätzbar bzw. – wie z.B.

bei Screening-Untersuchungen, bei denen Altersgruppen oftmals zugleich die Indikationsgruppen sind – berechenbar sind.

Das Modell dient dazu, die zwei Hauptdimensionen zu veranschaulichen, die zum Unterschied von Efficacy und Effektivität auf dem Niveau der Bevölkerung beitragen:

- ▲ Zugang zur Leistung (= Höhe der Rechtecke):
 - **systemseitige** Variablen wie das Vorhandensein einer Krankenversicherung, die Aufnahme der entsprechenden Technologie in den Leistungskatalog und räumliche Zugangsmöglichkeiten;
 - Variablen aufseiten der **Leistungserbringer**, insbesondere die Indikationsstellung zur Leistung, welche auch von systemseitigen Variablen wie der Vergütung beeinflusst werden kann;
 - **patientenseitige** Variablen wie empfunder Bedarf, Informationen etc.

- ▲ Durchführung der Leistung (= Breite der Rechtecke):
 - Variablen aufseiten der **Leistungserbringer**, insbesondere die technische Qualität der Leistungserbringung;
 - **patientenseitige** Variablen, insbesondere Compliance.

Letztere Dimension ist die bereits dargestellte klinische Betrachtungsweise, bei der nur die Patienten berücksichtigt werden, die tatsächlich die Maßnahme erhalten haben. In Abgrenzung von der bevölkerungsbezogenen Sichtweise sollte hier von „Klinischer Effektivität“ gesprochen werden.

Nun zu den Variablen der Leistungsdurchführung: Bezüglich der Patienten-Compliance liegen Daten vor, die bei der Medikamenteneinnahme unter Alltagsbedingungen deutliche höhere Diskontinuierungsraten als unter Studienbedingungen

belegen (s. Abb. 6.4). Im Modell zur Bevölkerungseffektivität führt dies zu einer Drehung der Efficacy-Linie, d.h. die größte Effektivitätseinbuße entsteht bei den Patienten mit dem größten erwarteten positiven Outcome (sofern diese ebenso häufig wie andere abbrechen). Umgekehrt verkleinert sich der erwartete gesundheitliche „Schaden“ für Patienten, die sich der Therapie unangemessenerweise unterziehen.

Auch zum Einfluss der Leistungsanbieter auf die Effektivität liegen etliche Daten vor. So berichten etwa Wennberg und Kollegen (1998) [20], dass die perioperativen Mortalitätsraten bei der Carotis-Endarteriektomie in Krankenhäusern, die an Studien teilnahmen, um 15% niedriger als in Nicht-Studien-Krankenhäusern mit vielen Eingriffen, um 25% niedriger als in Nicht-Studien-Krankenhäusern mit einer mittleren Anzahl von Eingriffen und 43% niedriger als in Nicht-Studien-

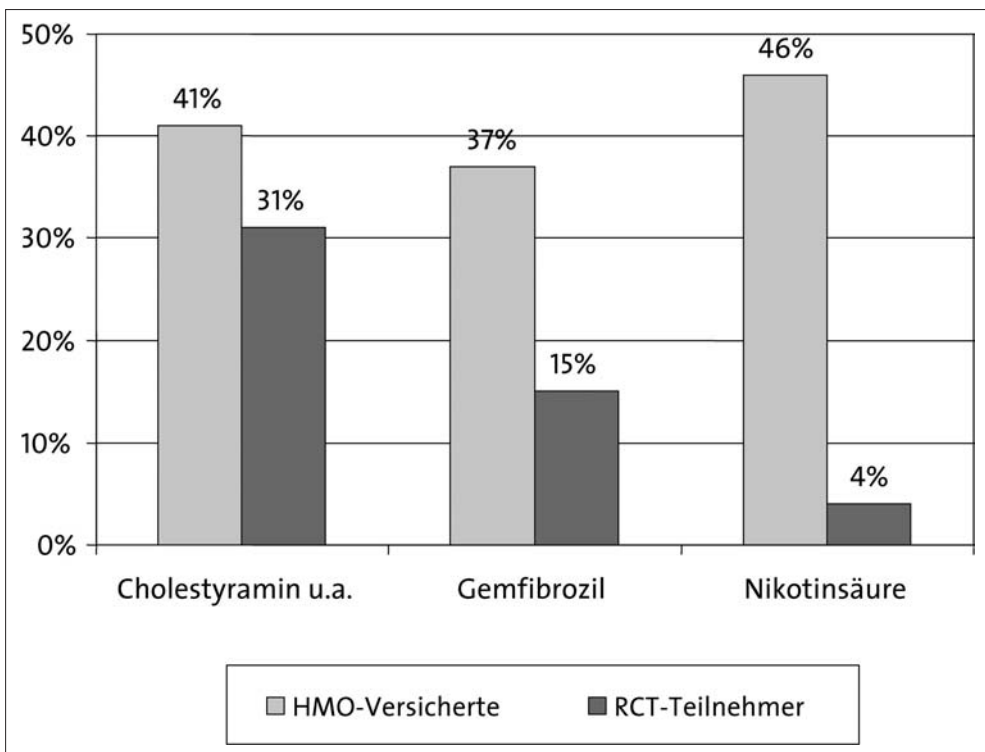


Abb. 6.4: Effektivität unter Alltagsbedingungen versus „Efficacy“: Diskontinuierung von Lipidsenkern als Beispiel für den Einfluss von Patientcompliance (eigene Darstellung nach [4])

Krankenhäusern mit geringer Eingriffsfrequenz waren.

Im Gegensatz zum patientenseitigen Einfluss führt die durch die Leistungserbringer verursachte verschlechterte technische Qualität zu einer Parallelverschiebung der Effektgröße gegenüber der Efficacy, d.h. auch Patienten mit einer Efficacy von 0 können von der Maßnahme eher Schaden als Nutzen erwarten. Als bevölkerungsmedizinischer Effekt ergibt sich: Das positive Outcome bei den angemessenen Indikationen wird geringer und das negative Outcome bei der unangemessenen Indikation größer. Selbstverständlich wirken die dargestellten Variablen in der tatsächlichen Versorgung nicht einzeln, sondern in Kombination. Dabei multiplizieren sich die auf den Zugang einwirkenden Effektminderungen mit denjenigen, die Unterschiede bei der Durchführung bedingen.

6.3 Ökonomische Daten für evidenzbasierte Systementscheidungen

Wir hatten eingangs bereits darauf hingewiesen, dass auch die benötigten (und vorhandenen) finanziellen Ressourcen bei der evidenzbasierten Gesundheitsversorgung eine Rolle spielen. Die Entscheidungskriterien hierzu können vereinfachend wiederum in zwei Anteile geteilt werden, nämlich in die Frage, ob der Effekt die Kosten rechtfertigt („Kosten-Effektivität“) und in die Kosten an sich, d.h. ob sich das System die Maßnahme leisten kann und will.

Es gibt einen gewissen Konflikt darüber, ob die ökonomischen Daten auf den Ergebnissen klinischer Studien oder auf routinemäßig gewonnenen Daten (z. B. von Krankenkassen) beruhen sollten. Erstere haben eine höhere interne Validität und letztere Daten eine höhere externe Validität und sind daher eher verallgemeinerbar. Im Prinzip haben beide Herangehensweisen ihre Berechtigung.

So kann eine ökonomische Studie zusammen mit einer klinischen Studie durchgeführt werden, wenn zusätzlich zu den klinischen Daten noch Daten zur Lebensqualität und zum Ressourcenverbrauch erhoben werden. Falls diese Herangehensweise gewählt wird, muss jedoch sichergestellt sein, dass der Ablauf innerhalb der Studie demjenigen unter Alltagsbedingungen zumindest ähnelt, dass relevante klinische Alternativen zum Vergleich mit einbezogen werden und dass die Untersuchungspopulation groß genug ist, um die ökonomischen Schätzungen auf eine solide Basis zu stellen. Trotz der Kombination mit einer klinischen Studie kann es noch zusätzlich notwendig sein, Ergebnisse über die Studienphase hinaus zu modellieren [21].

Die alternative Herangehensweise liegt darin, eine modellierende ökonomische Studie durchzuführen, bei der die Daten für die ökonomische Evaluation aus einer Vielzahl von Quellen gewonnen und für die ökonomische Evaluation synthetisiert werden. Diese Datenquellen könnten sowohl reine klinische oder ökonomische Studien sein, aber auch Datenbanken mit routinemäßig erhobenen Daten. Bei einer solchen Mischung von Datenquellen muss jedoch beachtet werden, dass die gewählte Herangehensweise transparent ist und die gemachten Annahmen dem Untersuchungsgegenstand angemessen sind.

Je nach Untersuchungsgegenstand und gewählter Herangehensweise beeinflussen eine Reihe von Faktoren die Verallgemeinerbarkeit und Übertragbarkeit der Ergebnisse von ökonomischen Evaluationen. Die mangelnde Übertragbarkeit betrifft dabei nicht nur die Übertragung der in einem Land gewonnenen Ergebnisse auf ein anderes Land, sondern kann auch innerhalb eines Landes gegeben sein. Insbesondere ist dies zu beachten, wenn die Kosteneffektivitäts-Evaluation auf Studienergebnissen und nicht auf unter Alltagsbedingungen gewonnenen Resultaten beruht. Da die Entscheidungsträger häufig

die an anderer Stelle gesammelten Ergebnisse für ihre eigene Entscheidung nutzen möchten, ist es wichtig, die Faktoren zu kennen, die die Verallgemeinerbarkeit bzw. Übertragbarkeit einschränken können.

Zu diesen Faktoren gehören insbesondere:

- ▲ die Epidemiologie der untersuchten Konditionen (z.B. Inzidenz und Prävalenz mit ihren Auswirkungen auf den positiven bzw. negativen Vorhersagewert von Untersuchungen),
- ▲ die Struktur des Gesundheitswesens (z.B. Aufgabenverteilung zwischen und innerhalb von Professionen, Finanzierungsmethoden für niedergelassene Ärzte und Krankenhäuser),
- ▲ klinische Entscheidungsfindung und Prioritätensetzung (arztseitige Faktoren) und
- ▲ patientenseitige Faktoren wie Bedarf, Nachfrage und Präferenzen.

6.4 Fallstudie Mammographie-Screening

Damit kommen wir zurück zu unserem Fallbeispiel, dem Mammographie-Screening. Für eine Entscheidung auf Systemebene werden auf jeden Fall Daten zur Inzidenz, Prävalenz und Mortalität der Erkrankung, der Bedeutung des Screenings im Verhältnis zur Therapie, die Evidenz der Wirksamkeit der vorgeschlagenen Intervention (in diesem Falle die Mammographie per ionisierender Strahlung) und Kostenschätzungen bzw. -hochrechnungen benötigt.

6.4.1 Bedeutung des Mamma-Karzinoms

Im Jahre 2004 verstarben in Deutschland insgesamt 435 000 Frauen. Das Mamma-Karzinom führte die Liste der karzinombedingten Todesursachen für Frauen mit 17 592 Toten (4,4% aller Todesfälle) an. Anders ausge-

drückt: Die Wahrscheinlichkeit für eine 60-jährige Frau, innerhalb der nächsten 10 Jahre an einem Mamma-Karzinom zu versterben, beträgt 0,7%, die Wahrscheinlichkeit, überhaupt zu versterben, liegt bei 12,6% [22]. Das Mamma-Karzinom trägt zur Gesamt mortalität also glücklicherweise nur in begrenztem Ausmaße bei. Anders verhält es sich mit der Inzidenz: Jährlich werden 48 000 Mamma-Karzinome diagnostiziert. Die altersspezifische Mortalität hingegen ist in Deutschland, wie in anderen Ländern auch, seit Jahren rückläufig. Bekannt ist auch, dass ein nicht unerheblicher Anteil von Frauen **mit**, aber nicht **an** einem Mamma-Karzinom verstirbt, wofür auch Zufallsbefunde bei Obduktionen sprechen.

Schon seit mehreren Jahren nimmt das Mamma-Karzinom einen hohen Stellenwert in der öffentlichen Berichterstattung und Diskussion ein. Seit vielen Jahren wird deshalb nicht nur von Frauengruppen, sondern auch von politischen Parteien und Einrichtungen wie der Europäischen Union die Einführung eines qualitätsgesicherten Mammographie-Screening-Programms gefordert. Länder wie die Niederlande, Großbritannien, Schweden oder Norwegen haben deshalb systematische Screening-Programme eingeführt. In der öffentlichen Wahrnehmung wird die Häufigkeit des Mamma-Karzinoms als Todesursache – wie auch die Effektivität der Mammographie – überschätzt [23].

6.4.2 Effektivität der Maßnahmen

Die Therapie des Mamma-Karzinoms wurde über die Jahre kontinuierlich weiterentwickelt, wobei erhebliche Fortschritte im Hinblick auf das Überleben in Form verbesserter Therapie nicht mehr verzeichnet werden konnten. Grundlegend geändert hat sich die Einstellung gegenüber der Radikalität der Therapie. Während in den Anfangsjahren möglich weitgehende chirurgische Eingriffe

zur Eradikation des Tumorherdes favorisiert wurden, zeigte sich in kontrollierten Studien, dass schon frühzeitig Tumorzellen in den Körper eindringen und somit eine systemische Krankheit vorliegt, deren Schicksal sich in der Regel nicht an der Stelle des Primärtumors entscheidet. Dies führte in der Folge zu wesentlich schonenderen Operationsverfahren. Hoffnungen konzentrieren sich auf die Früherkennung des Tumors. In 6 von 8 durchgeführten randomisierten kontrollierten Studien konnte gezeigt werden, dass mit der Mammographie eine statistisch signifikante Senkung der mammarkarzinombedingten Mortalität um ca. 25% möglich ist. Eine Senkung der Gesamtmortalität war aufgrund des relativ seltenen Ereignisses „Brustkrebstod“ trotz großer Fallzahlen nicht nachzuweisen. An der Validität der Studien wurden in späteren Jahren Zweifel geäußert – wegen fehlerhafter Randomisierung, ungleicher Gruppengrößen und nachträglicher Zuordnung von Todesfällen zur Kontrollgruppe [24].

Aufgrund des gewählten Studiendesigns, der Konsistenz des Ergebnisses über die Studien hinweg und der Übertragbarkeit der Ergebnisse auf das jeweilige Gesundheitssystem wird durch die damit befassten Einrichtungen wie der International Agency for Research on Cancer (IARC) oder der EU an der Wirksamkeit des Mammographie-Screenings für die Altersgruppe der 50–69-jährigen Frauen festgehalten. Ausschließlich für die Mammographie konnte eine solche Mortalitätssenkung nachgewiesen werden, gleichwohl hat sie ihre Limitationen. Nicht alle Karzinome kommen zur Darstellung (falsch negative Befunde), manche Strukturen imponieren wie Karzinome (falsch positive Befunde) und nicht jedes Karzinom beeinträchtigt tatsächlich das Leben der betroffenen Frau (Überdiagnose) [25].

Aus der Kenntnis der epidemiologischen Situation und der Ergebnisse der kontrollierten Studien lässt sich deshalb – die Einführung eines systematischen, qualitätsgesicher-

ten Mammographie-Screenings vorausgesetzt – annehmen, dass zwischen 3 000 und 5 000 brustkrebsbedingte Todesfälle in Deutschland potenziell vermieden werden könnten. Die Kombination von Effektivität der Intervention und Prävalenz der Erkrankung führt zu einer abwägenden Beurteilung des Mammographie-Screenings. Während 7 von 1 000 60-jährigen Frauen innerhalb von 10 Jahren an einem Mamma-Karzinom versterben, sind dies mit Mammographie-Screening 5–6. Die Wahrscheinlichkeit, dass die einzelne, am Screening teilnehmende Frau von der Maßnahme profitiert, ist – aufgrund der in Betracht aller Todesursachen eher geringen Mortalität des Mamma-Karzinoms – relativ gering. Soll eine entsprechende bevölkerungsbezogene Mortalitätsreduktion von 3 000–5 000 vermiedenen Karzinom-Toten erzielt werden, müssen mindestens ca. 3,5 Millionen Frauen jährlich an dieser Maßnahme teilnehmen. Wie bei Impfmaßnahmen auch, trügen deshalb sehr viele die Last, damit relativ wenige (gemessen an der Gesamtzahl der Teilnehmenden) profitieren.

6.4.3 Die Kosteneffektivität

Umfassende gesundheitsökonomische Berechnungen zur Kosteneffektivität des Mammographie-Screenings in Deutschland liegen – eigentlich überraschenderweise – nicht vor. Bei konservativ geschätztem direkten Aufwand von ca. 250 Mio. € jährlich (einschließlich der Kosten für Abklärung von falsch positiven Befunden, aber ohne indirekte Kosten durch Produktionsausfall) und ca. 4 000 vermiedenen Todesfällen ergeben sich Kosten von rund 60 000 € pro vermiedenem Todesfall. Bei einer Lebenszeitverlängerung um nur weitere drei Jahre entspräche dies 20 000 € pro gewonnenem Lebensjahr, ein im Vergleich zu internationalen Schwellenwerten für Kosteneffektivität durchaus akzeptabler Wert.

6.4.4 Präferenzen und Werte

Die Darstellung der Effektivität des Screenings aus bevölkerungsbezogener und individueller Sicht als Voraussetzung für eine informierte Entscheidung der einzelnen Frau ist dabei nicht unproblematisch. Je nach Sichtweise können die Vor- und Nachteile eines solchen Programms prononciert werden, wobei die reine Nebeneinanderstellung von Ergebnissen wenig aussagekräftig ist. Einmal mehr ist es deshalb eine gesellschaftliche und damit Werteentscheidung, wie über das Programm berichtet werden soll.

Auch die exakteste Darstellung der zugrunde liegenden Evidenz über Krankheitsgeschehen, die zur Verfügung stehenden diagnostischen und therapeutischen Optionen sowie die möglicherweise zu erzielenden Vorteile können in den seltensten Fällen eine solche Entscheidung vorwegnehmen. Im konkreten Fall des Mammographie-Screenings, das 10 Mio. Frauen der Altersgruppe von 50–69 betrifft, ist die Entscheidung Pro oder Contra daher eine Werteentscheidung. Gute Gründe lassen sich sowohl für als auch gegen ein solches Programm aufführen, was einem „Glas halb leer oder halb voll“-Dilemma entspricht. Sir Muir Gray, einer der Begründer der „Evidence-based health care“ spricht deshalb von einer „wertebasierten“ (value-based) Gesundheitsversorgung, deren Grundlage ein Wissens- und Informationsmanagement darstellt, bei dem das Konstrukt der Evidenz nach wie vor einen zentralen Stellenwert einnimmt.

Die Vorstellung, beispielsweise über HTA-Berichte (s. Kap. 16) eine ausschließlich sachbezogene, „technokratische“ Entscheidung treffen zu können, wird dieser Fragestellung nicht gerecht. Gleichwohl muss diskutiert werden, welches Maß an Information essenziell ist, um sich überhaupt für oder gegen eine solche Maßnahme auszusprechen. Wie Evidenz speziell für Entscheidungsträger aufbereitet werden sollte, ist inzwischen

selbst Forschungsgegenstand (vgl. etwa [26–28]).

Die parteienübergreifende Entscheidungsfindung des Deutschen Bundestags zur Einführung des Mammographie-Screenings im Jahre 2002 wurde detailliert bis hin zur Einrichtung eines aktiven Einladungssystems zur Ansprache aller Frauen in der Zielgruppe von 50 bis 69 Jahren getroffen und wird auch begründet. Bei der Darstellung der epidemiologischen Zahlenverhältnisse wird jedoch deutlich, dass abwägende Informationen, die auch Nachteile eines solchen Programms darstellen, offensichtlich nicht vorlagen. Hinzu kommt, dass gesundheitsökonomische Untersuchungen sowie eine versorgungsbezogene Gesundheitsberichterstattung im deutschen Gesundheitssystem noch nicht hinreichend verankert sind. Dies wäre gleichzeitig auch Voraussetzung dafür, eine „rationale“ Prioritätensetzung einzuleiten, bei der die eingesetzten Ressourcen im Hinblick auf ihre Nutzbringungs kritisch hinterfragt und entsprechende Weichenstellungen vorgenommen werden.

6.4.5 Gesetzgebungsinitiativen

Im Unterschied zu Leitlinien oder etwa der Anwendung der evidenzbasierten Entscheidungsfindung im klinischen Routinealltag haben Instrumente der evidenzbasierten Gesundheitsversorgung wie HTA-Berichte schnell Verbreitung gefunden. Der G-BA (s. Kap. 36) verwendet schon seit den 90er-Jahren HTA-Berichte für seine Entscheidungsfindung, mit der Einrichtung des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen im Jahre 2004 wurde die Entscheidungsunterstützung im G-BA im Sinne einer informierten Entscheidung weiter ausgebaut. Mit der Gesundheitsreform im Jahre 2007 soll auch die gesundheitsökonomische Komponente, die bislang noch nicht intensiv genutzt wurde, durch das IQWiG

bei der Beurteilung von Arzneimitteln aktiv aufgegriffen werden. Im Hinblick auf die Prioritätensetzung sind allenfalls Ansätze einer rationalen Priorisierung in einzelnen (Unter-)Ausschüssen des G-BA erkennbar.

Literatur

- [1] Hlatky MA et al., Comparison of predictions based on observational data with the results of randomized controlled clinical trials of coronary artery bypass surgery. *J Am Coll Cardiol.* (1988), 11, 237–245
- [2] Remington RD, Potential impact of exclusion criteria on results of hypertension trials. *Hypertension.* (1989), 13, I66–I68
- [3] Chalmers TC et al., Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. *N Engl J Med.* (1977), 297, 1091–1096
- [4] Sacks H, Chalmers TC, Smith H, Jr., Randomized versus historical controls for clinical trials. *Am J Med.* (1982), 72, 233–240
- [5] Colditz GA, Miller JN, Mosteller F, How study design affects outcomes in comparisons of therapy. I: Medical. *Stat Med.* (1989), 8, 441–454
- [6] Miller JN, Colditz GA, Mosteller F, How study design affects outcomes in comparisons of therapy. II: Surgical. *Stat Med.* (1989), 8, 455–466
- [7] Schulz KF et al., Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* (1995), 273, 408–412
- [8] Kleijnen J et al., (1997) So what's so special about randomisation? In: Maynard A, Chalmers I, Non-random reflections on health services research. On the 25th anniversary of Archie Cochrane's Effectiveness and Efficiency, 93–106. *BMJ Publ. Gr., London*
- [9] McKee M et al., Methods in health services research. Interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ.* (1999), 319, 312–315
- [10] Lewsey JD et al., Using routine data to complement and enhance the results of randomised controlled trials. *Health Technol Assess.* (2000), 4, 1–55
- [11] Fowler FJ, Jr. et al., Patient-reported complications and follow-up treatment after radical prostatectomy. *The National Medicare Experience: 1988-1990* (updated June 1993). *Urology.* (1993), 42, 622–629
- [12] Brook RH, Lohr KN, Efficacy, effectiveness, variations, and quality. Boundary-crossing research. *Med Care.* (1985), 23, 710–722
- [13] Tugwell P et al. (1984) A framework for the evaluation of technology: the technology assessment iterative loop. In: Feeny D, Guyatt G, Tugwell P, The Canadian Medical Association, The Institute for Research on Public Policy, Health care technology: Effectiveness, efficiency and public policy, 41–56.
- [14] Busse R (2006) Leistungsmanagement im Gesundheitswesen – Einführung und methodische Grundlagen. In: Busse R, Schreyögg J, Gericke C, Management im Gesundheitswesen, 12–23. Springer, Berlin
- [15] Herrin J et al., Effect of panel composition on physician ratings of appropriateness of abdominal aortic aneurysm surgery: elucidating differences between multispecialty panel results and specialty society recommendations. *Health Policy.* (1997), 42, 67–81
- [16] Wennberg JE, The paradox of appropriate care. *JAMA.* (1987), 258, 2568–2569
- [17] Yusuf S et al., Effect of coronary artery bypass graft surgery on survival: overview of 10-year results from randomised trials by the Coronary Artery Bypass Graft Surgery Trialists Collaboration. *Lancet.* (1994), 344, 563–570
- [18] Smith GD, Egger M, Incommunicable knowledge? Interpreting and applying the results of clinical trials and meta-analyses. *J Clin Epidemiol.* (1998), 51, 289–295
- [19] Andrade SE et al., Discontinuation of anti-hyperlipidemic drugs – do rates reported in clinical trials reflect rates in primary care settings? *N Engl J Med.* (1995), 332, 1125–1131
- [20] Wennberg DE et al., Variation in carotid endarterectomy mortality in the Medicare population: trial hospitals, volume, and patient characteristics. *JAMA.* (1998), 279, 1278–1281
- [21] Busse R, Gesundheitsökonomie – Ziele, Methodik und Relevanz. *Bundesgesundheitsbl Gesundheitsforsch Gesundheitschutz.* (2006), 49, 3–10

- [22] Fletcher SW, Elmore JG, Clinical practice. Mammographic screening for breast cancer. *N Engl J Med.* (2003), 348, 1672–1680
- [23] Chamot E, Perneger TV, Misconceptions about efficacy of mammography screening: a public health dilemma. *J Epidemiol Community Health.* (2001), 55, 799–803
- [24] Gotzsche PC, Olsen O, Is screening for breast cancer with mammography justifiable? *Lancet.* (2000), 355, 129–134
- [25] Gotzsche PC, Nielsen M, Screening for breast cancer with mammography. *Cochrane Database Syst Rev.* (2006), CD001877
- [26] Lavis JN et al., Use of research to inform public policymaking. *Lancet.* (2004), 364, 1615–1621
- [27] Sheldon TA, Making evidence synthesis more useful for management and policymaking. *J Health Serv Res Policy.* (2005), 10 Suppl 1, 1–5
- [28] Mays N, Pope C, Popay J, Systematically reviewing qualitative and quantitative evidence to inform management and policymaking in the health field. *J Health Serv Res Policy.* (2005), 10 Suppl 1, 6–20

7 Pluralismus in der Evaluation

Karl Wegscheider

7.1 Quellen der Evidenz

7.1.1 Die Rolle der Wissenschaftsmethodik in der Evidenzbasierten Medizin (EbM)

In den folgenden Kapiteln dieses Buches werden verschiedene Techniken zur Evidenzfindung und -bewertung und die dahinter stehenden konzeptionellen Vorstellungen ausführlich für viele Gebiete und Anwendungsbeispiele dargestellt. Bevor wir uns im Folgenden auf diesen teilweise mühsamen, insgesamt aber doch erfolversprechenden Weg zur rationalen Basierung ärztlichen Handelns begeben, sollten wir jedoch einen Moment innehalten und uns der Frage zuwenden: Wie entsteht eigentlich die Evidenz, die wir als EbM-Praktizierende aufzuspüren und umzusetzen versuchen? Für die EbM-Gründerväter war das keine drängende Frage. Sie hatten im Rahmen ihrer Forschungstätigkeit mehr neue Evidenz erzeugt, als die Praktiker ihrer Zeit umzusetzen bereit und in der Lage waren. Zudem war ihnen als Methodikern der Prozess der Evidenzgenerierung wohl vertraut. Dem EbM-Praktiker unserer Tage, der die Erfordernis der Evidenzbasierung zum Maßstab seines täglichen Handelns gemacht hat, stellt sich die Situation möglicherweise anders dar. Immer häufiger trifft er bei seinen Recherchen auf Fallkonstellationen, für die überzeugende Evidenz nicht leicht verfügbar ist (wie etwa für weite Teile der Chirurgie oder für die Homöopathie), bei denen ein Konflikt zwischen verschiedenen Evidenzaussagen zu bestehen scheint (zum Beispiel bei der Behandlung

multimorbider Patienten) oder gar die Fachwelt in tiefen Konflikten über die Interpretation von Studien und Meta-Analysen verstrickt ist (wie etwa beim Mammographie-Screening, [1; 2]). In diesen Fällen führen häufig auch die erlernten Grundtechniken wie Literaturrecherche und Evidenzbewertung anhand von Checklisten nicht zum Ziel. Es erscheint daher selbst für altgediente Praktiker der Evidenzbasierten Medizin unumgänglich, sich mit den wissenschaftlichen Ansätzen zur Evidenzgewinnung, das heißt mit der Methodik der Therapie-Evaluation, auseinanderzusetzen und sich hierzu ein eigenes Urteil zu bilden.

7.1.2 Evidenzgenerierung – eine Aufgabe der Wissenschaft

Kann es Evidenz nicht auch ohne Wissenschaft geben? – Wird im EbM-Begriffsrahmen von „Evidenz“ gesprochen, so ist damit grundsätzlich empirisch begründete Erkenntnis gemeint, die durch systematische Erprobung und Beobachtung von therapeutischen Prozessen entsteht, d.h. durch Therapie-Evaluation mit wissenschaftlichen Methoden. Die EbM-Gründerväter grenzen die Evidenz ausdrücklich von klinischer Expertise, Intuition, unsystematischer Beobachtung oder rein pathophysiologischer Begründung ab. Das Grundlagenwissen und die Soft Skills des Arztes spielen zwar eine große Rolle im therapeutischen Prozess. Bei der Bewertung der Wirksamkeit und des therapeutischen Nutzens einer Therapie (und nur dort) sind sie jedoch nachrangig hinter der aus empiri-

scher Forschung gewonnenen Erkenntnis einzuordnen, soll der Anspruch der Evidenzbasierung erhoben werden [3]. Die Aufgabe der Evidenzgenerierung wird damit der Wissenschaft und der dort vorherrschenden Methodenlehre überantwortet, die hierbei auf eine lange Tradition bauen kann, die wesentlich weiter zurückreicht als die EbM-Bewegung.

7.1.3 Die Evidenzbasierung der Medizin – der Weg zur perfekten Behandlung?

Welche Erwartungen darf man realistischerweise an die von der Wissenschaft zur Verfügung gestellten Ergebnisse haben? In welchem Umfang wird die wissenschaftlich begründete Evidenz die praktische Medizin durchdringen können? Werden wir eines Tages alle therapeutischen Entscheidungen evidenzbasiert fällen können?

In der EbM-Ausbildung werden die Techniken der Evidenzfindung und -bewertung systematisch gelehrt. Wichtiges Hilfsmittel sind dabei die verbreiteten Evidenzklassensysteme (vgl. Kap. 12, Tab. 12.1). Auf der höchsten Stufe befinden sich systematische Reviews und Meta-Analysen randomisierter Studien. Für verschiedene medizinische Gebiete steht eine stetig wachsende Zahl solcher Reviews zur Verfügung. Damit scheint der Weg zur perfekten Medizin vorgezeichnet zu sein: Muss uns nicht lediglich Forscherfleiß zu jeder therapeutischen Frage die auf randomisierten Studien aufbauende ultimative Meta-Analyse zur Verfügung stellen, aus der Evidenz dann quasi automatisch entsteht? Gut, das wird vielleicht in manchen medizinischen Gebieten noch etwas dauern. Zwischenzeitlich müssen wir uns für therapeutische Entscheidungen an den Studien der jeweils aktuell höchsten Evidenzstufe orientieren. Bringt aber hier Recherche fleiß nicht im Grunde auch schon die richtige evidenzbasierte Entscheidung hervor? Sind wir

nicht eigentlich schon im Gelobten Land der Evidenzbasierung angekommen, in dem die Evidenz wie Manna vom Himmel fällt und im Überfluss vorhanden ist, dem Land, in dem die Milch der Erkenntnis und der süße Honig der rational begründeten Gesundheit fließen und die perfekte Therapie nur noch eine Frage der Umsetzung ist? Müssen wir, um die letzten Lücken zu schließen, die Evidenzbasierte Medizin nicht einfach nur angehen wie etwa das Humane Genome Project zur Entschlüsselung des menschlichen Gens oder das Manhattan-Projekt zum Bau der Atombombe, und der Weg über die perfekte Therapie zum gesunden Leben ist uns praktisch sicher?

Verfolgt man die aktuellen Diskussionen zur Umgestaltung des Gesundheitssystems, so scheint es, dass sich die skizzierte Heilerwartung an die Evidenzbasierte Medizin in vielen Köpfen festgesetzt hat, obwohl sie den EbM-Gründervätern in ihrem kritischen Wissenschaftsverständnis völlig fern gelegen hätte und obwohl sie zu keinem Zeitpunkt zum expliziten Kanon der EbM-Bewegung gehört hat. Evidenzbasierung scheint manchem der Schlüssel zu jedwedem Erfolg im Gesundheitswesen zu sein, eine Art Zauberstab, von dem man sich alles erhofft: von der Sanierung der Haushalte bis zur Austreibung der medizinischen Teufel wie Quacksalberei und Naturheilkunde. In einer säkularisierten Welt besteht Bedarf nach Gottesersatz, da bietet sich die positive Utopie einer rational begründeten Gesundheit von System und Patient, wie sie die Evidenzbasierte Medizin zu verheißen scheint, dem aufgeklärten Geist geradezu an. Dabei ist diese Utopie nichts anderes als eine Fortschreibung des mechanistischen Weltbildes des 19. Jahrhunderts, das sich die Welt wie ein Uhrwerk vorstellte, in dem man nur an den richtigen Schraubchen drehen müsse, um gewünschte Ergebnisse zu erzielen. Mit gemeinsamer Anstrengung erscheint dann alles machbar. In dieser deterministisch geprägten Weltsicht

ist entsprechend für Nachdenklichkeit und Zaudern bei der Evidenzfindung kein Platz. Beides steht im Verdacht, fortschritts- und innovationsfeindlich zu sein. Ebenso schädlich wäre dann auch ein Pluralismus in der Therapie-Evaluation, der Sand ins Getriebe dieses Uhrwerks streuen würde.

7.1.4 Der Wissenschaftsprozess – eine Spirale ohne Ende

Würde diese Utopie tragen – wer könnte ernsthaft dagegen sein, diesen Weg konsequent zu verfolgen? Allerdings: die gesamte Utopie basiert auf der Vorstellung, dass die Wissenschaft Evidenz in beliebigem Umfang mit einem allgemein anerkannten etablierten Verfahren erzeugen könne. Die Aufgabe wäre dann rein quantitativ und könnte durch simple Investition von Forschungsgeldern mit Erfolgsgarantie vollzogen werden. Leider basiert diese Vorstellung aber auf einem verkürzten und im Kern falschen, naiven Verständnis von Wissenschaft, das von den tatsächlichen Möglichkeiten wie von den Begrenzungen und Versuchungen von Wissenschaft und Wissenschaftsprozess nichts weiß. Tatsächlich ist das Ringen um Evidenz ein mühseliger, fehleranfälliger, langwieriger Prozess ohne definitives Ende, eine Herausforderung und ein Wagnis, begleitet von Täuschung, Selbsttäuschung, Missverständnissen und Trugschlüssen selbst bei ehrbaren Wissenschaftlern. Ein Königsweg zur Erkenntnis eröffnet sich nicht und steht auch nicht in Aussicht. Mit jeder neuen Erkenntnis stößt Wissenschaft in neue komplexere Zusammenhänge vor. Mit jeder neuen Antwort vergrößert sich die Zahl der Fragen, und zwar nicht nur die Zahl der möglicherweise weniger wichtigen Detailfragen. Vielmehr werden gerade die scheinbar unerschütterlichen Grundfesten unseres Verständnisses von klinischen Abläufen regelmäßig neuen kritischen Betrachtungen aus-

gesetzt. Ständig werden an grundlegenden Vorstellungen in schneller Folge größere und kleinere Korrekturen erforderlich. Die Geschwindigkeit des Wandels scheint mehr durch die Begrenztheiten unseres Bewusstseins und unsere soziale Organisation gesteuert zu sein als durch die objektive Erfordernis. Natürlich gibt es dabei auch zyklische Phänomene (Moden), die ein starkes Individuum abkürzen könnte. Insgesamt ist jedoch die schnelle Adaption und die flexible Einstellung auf neue Gegebenheiten nicht etwa eine Begrenzung, sondern ein Qualitätsmerkmal guter Wissenschaft. Der an der Forschungsfront tätige Wissenschaftler erlebt deshalb Evidenz als etwas in ständiger Wandlung Befindliches, ein Moving Target, das beobachtet und nachverfolgt, aber wie ein wildes Tier nur fixiert werden kann, indem man ihm das Leben nimmt. Wissenschaft kann deshalb von Natur aus die Erwartungen des „users“, der die Ergebnisse wie ein technisches Gerät einsetzen will und reibungslos funktionieren bis zum Ende der Garantiezeit erwartet, nicht erfüllen. Wer sich in seiner praktischen Arbeit auf wissenschaftliche Ergebnisse stützen will, muss sich, um dem Anspruch von Wissenschaftlichkeit zu genügen, auf die Ambivalenzen wissenschaftlicher Bewertungen und Schlussfolgerungen einlassen. Anderenfalls läuft er Gefahr, zum Opfer von Zufälligkeiten oder systematischen Fehlern zu werden.

7.2 Die Methodik der Evidenzgenerierung

7.2.1 Die randomisierte Studie – zu Recht der Goldstandard?

Die Gerüchte wollten einfach nicht verstummen: Das verbreitete Medikament XY, erste Therapieoption in der postoperativen Versorgung auf der Intensivstation, sollte Venenthrombosen verursachen! Eine Reihe

von Spontanberichten stützte den Verdacht, aber diese könnten Folge der medialen Aufmerksamkeit sein. Wie auch immer, die Umsätze gingen zurück, und so wurde die renommierte Universitätsklinik beauftragt, auf der Basis der Daten in ihrer pharmako-epidemiologischen Datenbank den Verdacht zu untersuchen. Tabelle 7.1 zeigt das Ergebnis. Von 1507 verzeichneten Nutzern des Medikaments hatten 1,1% Venenthrombosen auf der Intensivstation erfahren, von den Nichtnutzern 1,3% (Odds-Ratio=0,85, 95%-CI:

Tab. 7.1: Postoperative Venenthrombosen in Abhängigkeit von der Einnahme des Medikamentes XY

Medika- ment XY	Venenthrombose		Σ
	nein	ja	
nein	1410 (98,7%)	19 (1,3%)	1429 (100%)
ja	1490 (98,9%)	17 (1,1%)	1507 (100%)
Σ	2900 (98,2%)	36 (1,2%)	2936 (100%)

Tab. 7.2: Postoperative Venenthrombosen in Abhängigkeit von der Einnahme des Medikamentes XY, nach Geschlecht

Medika- ment XY	Venenthrombose		Σ
	nein	ja	
Männer			
nein	90 (100%)	0 (0,0%)	90 (100%)
ja	970 (99,4%)	6 (0,6%)	976 (100%)
Σ	1060 (99,4%)	6 (0,6%)	1066 (100%)
Frauen			
nein	1320 (98,6%)	19 (1,4%)	1339 (100%)
ja	520 (97,9%)	11 (2,1%)	531 (100%)
Σ	1840 (98,4%)	30 (1,6%)	1870 (100%)

0,42–1,69). Das Medikament wies sogar einen schwachen Trend zur Vermeidung von Thrombosen auf, angesichts der Breite des Konfidenzbereiches verbot es sich allerdings, hier von einer protektiven Wirkung überhaupt nur zu sprechen. Immerhin: das Medikament schien entlastet.

Nach einem Jahr erschien ein grundlegender Aufsatz zur Rolle der Hormone bei der Entstehung von Venenthrombosen. Daraufhin analysierte die Universitätsklinik die Daten noch einmal getrennt für Männer und Frauen in der Erwartung, dass möglicherweise nur bei den Frauen oder nur bei den Männern ein Effekt zu sehen sei, der von der anderen Gruppe jeweils kompensiert würde. Das Ergebnis zeigt Tabelle 7.2. Überraschend zeigte sich jetzt in beiden Gruppen eine klare Tendenz zu einer Risikoerhöhung unter Medikamenteneinnahme.

Wie ist das möglich? In beiden Subgruppen gibt es bei unterschiedlichem Basisrisiko einen gleichlautenden Trend, und trotzdem weist die Gesamttafel ein neutrales Ergebnis auf? Die Erklärung liegt in den unterschiedlichen Verschreibungsgewohnheiten bei Männern und Frauen. Offensichtlich wurde das Medikament wesentlich häufiger bei Männern als bei Frauen verordnet. Dadurch weisen in der Gesamttafel die Zeilen eine unterschiedliche Geschlechterzusammensetzung auf. Der Medikamenteneffekt war mit dem Geschlechtseffekt vermengt („konfundiert“). Dadurch wurde der Effekt kompensiert. Es hätte auch umgekehrt kommen können: Hätten nicht die Männer, sondern die Frauen ein erhöhtes Basisrisiko gehabt, dann hätte die Gesamttafel einen starken Medikamenteneffekt aufgewiesen, selbst wenn in den Untergruppen kein klinisch relevanter Effekt vorhanden gewesen wäre. Der scheinbare Widerspruch wird auch „Simpsons Paradox“ genannt nach dem englischen Mathematiker, der ihn als erster beschrieben hat [4].

Wer je an einem wissenschaftlichen Projekt beteiligt war, bei dem im Nachhinein

eine das Ergebnis verändernde Einflussgröße (engl. „Confounder“) entdeckt wurde, die die medizinische Schlussfolgerung ins Gegenteil wendet, dem sitzt der Schreck tief in den Gliedern. Dabei liegt entgegen landläufiger Meinung das Problem nicht in den bekannten und längst identifizierten Confoundern („Overt Bias“, [5]). Diese lassen sich nämlich mit statistischen Methoden kontrollieren. Das Problem sind die Confounder, die bisher der Aufmerksamkeit entgangen sind, oder die bekannt sind, aber leider nicht gemessen werden können. Sie können eine dem Auge verborgene Verfälschung erzeugen, einen „hidden bias“ [5]. Der Hidden Bias ist das Damokles-Schwert der epidemiologischen Forschung. Die Möglichkeit der Existenz eines übersehenen Confounders gibt dem Ergebnis jeder Beobachtungsstudie etwas grundsätzlich Vorläufiges.

Es gibt nur eine Möglichkeit, das Problem des Hidden Bias unter Kontrolle zu bekommen: Man muss die Assoziationen zwischen der Therapie-Entscheidung und (allen!) möglichen Confoundern zerstören. Dazu muss man im Rahmen einer Studie die Therapien den Patienten zufällig zuteilen bzw. äquivalent die Patienten den Therapien zuweisen (randomisieren). Damit wird zwar Confounding nicht sicher vermieden, aber den Gesetzen der Wahrscheinlichkeitsrechnung unterworfen. Damit lässt sich das Problem des Confounding mit den üblichen Techniken zur Beherrschung zufälliger Effekte (Konfidenzbereiche, p-Werte) kontrollieren.

Wegen des Confounderproblems (und nur wegen des Confounderproblems) stehen randomisierte klinische Studien (RCTs) an vorderster Stelle der Evidenzhierarchie. Diese Einordnung steht und fällt damit, inwieweit man das Confounderproblem für das wichtigste Problem der Evidenzgewinnung hält. International besteht unter Methodikern seit 50 Jahren weitgehend Einigkeit, dass dem so ist. Insbesondere in der Medikamentenzulas-

sung erscheint die Idee kaum erträglich, dass ein Medikament dank eines Huckepackeffektes durch einen Confounder auf den Markt kommt, sei es durch vorgetäuschte Wirksamkeit oder durch ein verdecktes Sicherheitsproblem.

Allerdings: für die Beherrschung des Hidden Bias ist ein Preis zu zahlen. Dieser Preis besteht in den Einschränkungen, die zur Realisierung des Studienkonzeptes erforderlich sind. Nicht jeder Patient, der später das Medikament bekommen wird, wird bereit sein, an einer randomisierten Studie teilzunehmen. Praktische, juristische, ethische und ökonomische Gründe werden die Studienorganisatoren zwingen, enge Ein- und Ausschlusskriterien zu definieren. Damit steht die Generalisierbarkeit der Studienergebnisse grundsätzlich in Frage [6].

Des Weiteren ist auch die Generalisierbarkeit nach innen keineswegs selbstverständlich. Die randomisierte Studie geht von der Fiktion eines homogenen Patientenkollektivs aus. Was ist, wenn es Untergruppen mit abweichenden Effekten gibt, so genannten Interaktionseffekten (nicht mit dem oben erwähnten Confounding zu verwechseln)? Prinzipiell können solche Effekte mit Interaktionstests (nicht mit Subgruppenanalysen) identifiziert werden, aber meistens reicht die Power der Studien hierzu nicht aus.

Auch randomisierte Studien können im Übrigen trotz erfolgreicher Confounderkontrolle auch anderen Arten von Verzerrungen unterliegen (s. Abb. 7.1, Spalte RCT). Wenn etwa Patienten, Ärzte oder Beurteiler nicht verblindet werden können, so kann es zu verzerrenden Beobachtungsfehlern kommen. Es kann eine unterschiedliche Compliance in den Behandlungsgruppen geben – mit dem möglicherweise verfälschenden Effekt, dass in beiden Gruppen unterschiedlich viele Patientenangaben fehlen (Differential Missings). Der Outcome-Parameter kann unter Effektivitätsgesichtspunkten so gewählt

Quelle	Meta-analyse Review	RCT	Kohortenstudie	Fall-/Kontr.-studie	Register	Einzelfall-Beurteilung
Overt Bias	●	-	-	-	-	●
Hidden Bias	●	-	●	●	●	●
Zufall	>5%	5%	5%	5%	5%	nicht bezifferbar
Mangelnde Verblindung	●	(●)	●	●	●	●
Differential missings	●	●	●	●	●	●
Mangelnde Repräs.	●	●	(●)	●	-	(●)
Heterogenität	(●)	●	(●)	●	-	●
Ungeeigneter Outcome	●	●	●	●	●	●

Abb. 7.1: Potenzielle Trugschlussquellen nach Studientypen. Markiert sind diejenigen Störeffekte, die beim jeweiligen Studientyp nicht kontrolliert (schwarzer Sprengsatz) bzw. nur unvollständig kontrolliert (grauer Sprengsatz in Klammern) werden.

sein, dass er klinisch fragwürdig ist. Und es können bei der Auswertung zusätzliche Vergleiche angestellt werden, die durch die Randomisierung nicht vor Confounding geschützt sind (etwa Vorher-Nachher-Vergleiche), die durch die Überschrift „Randomisierte Studie“ aber unangemessen aufgewertet werden.

Die Begrenzungen der randomisierten Studien zeigen sich insbesondere bei Sicherheitsproblemen. Eine verlässliche Bewertung der Sicherheitsrisiken ist zum Zeitpunkt der Zulassung auf der Basis der vorliegenden randomisierten Studien, d.h. der quantitativ begrenzten Erfahrung aus idealisierten Studienpopulationen, in der Regel nicht möglich.

7.2.2 Beobachtungsstudien: Repräsentativität vs. Verzerrung

Entfällt die Randomisation, so kann die Studienpopulation in der Regel breiter angelegt werden. Den Patienten werden die Therapien nicht mehr vom Studiendesign vorgeschrieben, es werden lediglich die Ergebnisse

beobachtet. In geplanten epidemiologischen Studien ist man dennoch nicht völlig frei in der Wahl der Studienpopulation. Einschlusskriterien müssen eingehalten werden, um Verzerrungen zu minimieren, und die Patienten müssen einer Langzeitbeobachtung zustimmen. Nur systematische Register sind, wenn sie das Kriterium der Vollständigkeit einhalten, wirklich repräsentativ für das Geschehen im medizinischen Alltag, allerdings um den Preis zusätzlicher Verzerrungsquellen.

Prinzipiell lässt sich das Problem des Overt Bias, d.h. der bekannten Confounder, durch entsprechend angelegte Studiendesigns (Stratifizierung, Matching) bzw. durch – allerdings aufwendige – statistische Analysemethoden beherrschen. Der Hidden Bias und zahlreiche weitere, auch bei randomisierten Studien mögliche Verfälschungsquellen müssen allerdings in Kauf genommen werden (vgl. Abb. 7.1, Spalten 4, 5 und 6). Letztlich „tauscht“ man beim Übergang von der randomisierten zur nicht-randomisierten Vergleichsstudie eine größere Repräsentativität der Aussagen gegen das erhöhte Risiko ei-

nes Hidden Bias ein – ein Handel, auf den sich viele Wissenschaftler nicht einlassen würden.

7.2.3 Systematische Reviews/ Meta-Analysen: ein methodischer Fortschritt?

Wenn denn die einzelnen Studien aus ihren methodischen Begrenzungen nicht herausfinden: können wir dann nicht dem Bewertungsdilemma entkommen, indem wir mit einem geschickten Mechanismus die qualitätsgesicherte Information aus den Studien herausfiltern und zu einer endgültigen Aussage verdichten?

Die Idee erscheint plausibel. Die inzwischen zahlreich verfügbaren Meta-Analysen und systematischen Reviews folgen dieser Leitvorstellung. Sie können aber natürlich nicht grundsätzlich besser sein als die verwendeten Einzelstudien. Dieses Problem ist den Verfassern durchaus bewusst. Bei der Auswahl der verwendeten Studien werden deshalb strenge Qualitätsmaßstäbe angelegt. Im Laufe der Jahre hat sich entsprechend eine eigene wissenschaftliche Methodik zur Erstellung der Reviews herausgebildet, die kollektiv gehütet wird. Die Reviews aus kontrollierten klinischen Studien stehen inzwischen an erster Stelle in den Evidenzklassensystemen. Schaffen sie die ersehnte höchste Evidenz?

Ein Grunddilemma begleitet diese Arbeit: Je strenger man die Qualitätskriterien festlegt, desto schmaler wird die empirische Basis und desto zweifelhafter wird die Repräsentativität. Auf der Ebene der Studienausswahl wiederholt sich das Problem, das sich auf der Ebene der Patientenauswahl bei einzelnen Studien stellte. Man muss sich entscheiden zwischen einer hohen Qualität des Vergleichs Therapie/Kontrolle und weitgefaster Generalisierbarkeit der Ergebnisse. Beide Ziele sind selten simultan zu erreichen.

Hinzu kommt: die Durchführung von Studien folgt keinem zentralen Entwicklungsplan eines alles dirigierenden Studienleiters, sondern ist das Resultat komplizierter interessengesteuerter gesellschaftlicher Aushandlungsprozesse. Sie führen häufig dazu, dass Therapien, mit denen viel Geld verdient werden kann, in der Studienlandschaft überrepräsentiert sind („Blockbuster-Bias“), während andererseits Patientengruppen wie Kindern, Multimorbiden oder Alten der Zugang zu den in Meta-Analysen bevorzugten RCTs praktisch verwehrt ist. Diese Fokusverschiebung können Meta-Analysen nicht ausgleichen.

Aus methodischer Sicht stellen somit die Reviews ausschließlich einen quantitativen Fortschritt dar. In der methodischen Qualität sind die Reviews einzelnen Studien eher unterlegen, weil sie im Tausch gegen einen Fortschritt bei der statistischen Power und größere Vollständigkeit des Überblicks verstärkt Problemen mit der Qualität der Studien und der Heterogenität der Effekte ausgesetzt sind.

7.2.4 Einzelfallstudien/Heilversuche: die fehlende Kontrolle

Auch in einer idealen Welt, in der alle Studien ohne Qualitätsmängel durchgeführt werden, kein Hidden Bias in den Beobachtungsstudien auftritt, die randomisierten Studien generalisierbar sind und die Meta-Analysen keine Heterogenität erkennen lassen, bleibt eine Frage offen: Wird die nach besten Studierkenntnissen gewählte Therapie beim aktuellen Patienten helfen? Diese Frage lässt sich in der Tat mithilfe statistischer Methoden nicht beantworten. Genaue Aussagen lassen sich aus Studien nur für große Patientengruppen ableiten. Je genauer man nach der Risikofaktoren-Konstellation eines einzelnen Patienten fragt, desto ungenauer wird die statistische Aussage, die man aus Studi-

endaten gewinnen kann. Sollte ich da nicht gleich die gewaltigen Mühen, die mit der Durchführung großer Studien einhergehen, vermeiden und mich ganz auf die Bewertung der Wirksamkeit im Einzelfall beschränken, Heilversuche unternehmen, ganzheitliche Bewertungen unter Berücksichtigung der vollständigen Patientengeschichte vornehmen, darüber berichten und anderen die passenden Analogieschlüsse bei ihren Patienten überlassen? Weiß nicht der einfühlsame Arzt tief im Inneren, ob er dem Patienten geholfen hat oder nicht [7]?

Wie verlässlich ist das Wirksamkeitserlebnis des Arztes? Abb. 7.2 zeigt eine Interpretationsaufgabe aus dem Methoden-Unterricht, die auf echten Praxisdaten beruht. Es handelt sich um die Blutdruckverläufe von Patienten, die vormittags zur jeweils gleichen Zeit mit systolischen Blutdruckwerten über 150 mm Hg in ihrer Arztpraxis erschienen sind. Im Unterricht werden Kursteilnehmer nun gebeten, den Therapieerfolg aus den dargestellten

Kurven zu bewerten. Fast alle konstatieren schnelle Wirksamkeit und Nachhaltigkeit der Therapie, die Kausalität wird auch auf Nachfrage nicht in Zweifel gezogen.

Tatsächlich waren die Patienten unbehandelt. Für die Darstellung wurden aus einem durchschnittlichen Praxiskollektiv ohne Blutdrucktherapie (Abb. 7.3) lediglich die Patienten ausgewählt, die um 11 Uhr einen erhöhten Blutdruckwert aufwiesen. Aufgrund der hohen Spontanvariabilität des Blutdrucks wählt man durch diese Vorgehensweise bevorzugt Patienten aus, die gerade zufällig erhöhte Werte aufweisen. Sie kehren nachfolgend zu ihrem eigenen Mittelwert zurück und verbleiben dort: „They regress towards the mean“, und so heißt auch der Trugschluss, der hier vorliegt.

Im Einzelfall hat ein bewertender Arzt keine Möglichkeit, eine Medikamentenwirkung vom Spontanverlauf zu unterscheiden. Aus psychologischer Forschung wissen wir, dass Ärzte wie andere Menschen eine große

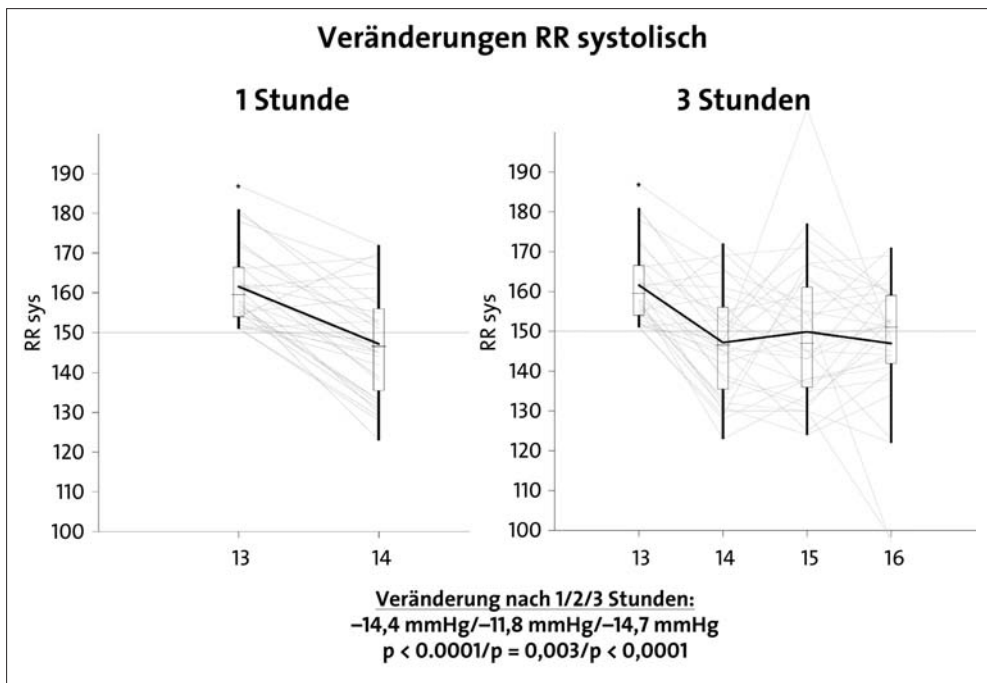


Abb. 7.2: Interpretationsaufgabe: Belegen die gezeigten Blutdruckprofile von 32 Praxispatienten die Wirksamkeit einer Intervention?

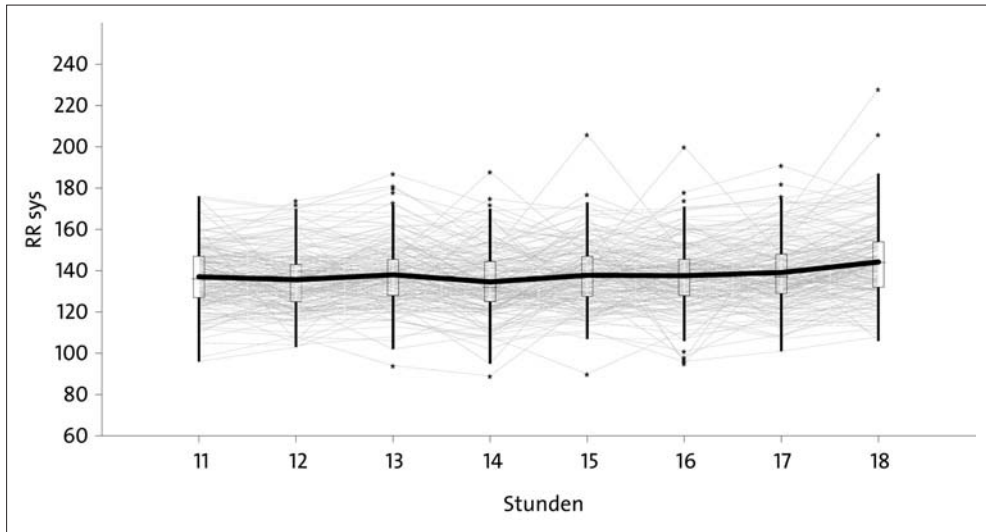


Abb. 7.3: Auswahlkollektiv zur Interpretationsaufgabe aus Abb. 7.2: Boxplot-gebündelte Blutdruckprofile 185 unbehandelter Routinepatienten aus verschiedenen Arztpraxen, erhoben im Rahmen von 24h-Messungen in der Screening-/Pilotphase einer randomisierten Studie.

Bereitschaft haben, zufällige Verbesserungen auf dem Konto „Therapieerfolg“ zu verbuchen. Jenseits aller Studien dürfte die bekannte Spontanvariation Quelle manchen Heilungserlebens in der täglichen ärztlichen Praxis sein. Dieser Trugschluss und unsere Begabung zur selektiven Wahrnehmung sind im Übrigen die eigentliche Ursache dafür, dass die Effekte in kontrollierten Studien fast immer geringer ausfallen als vorher angenommen. Nicht ärztliche Erfahrung, sondern **ausschließlich** das Mitführen einer Kontrollgruppe kann den Arzt vor dem vorgeführten Trugschluss in der Therapiebewertung bewahren und ihn ebenso wie die wissenschaftliche Gemeinschaft in die Lage versetzen, zwischen Spontanvariation und Therapieerfolg zu unterscheiden. Hauptsächlich deswegen (und nicht aus allgemeiner erkenntnistheoretischer Überlegung) sind kontrollierte Studien unerlässlich. Bei der Einzelfallbewertung fehlt die Kontrolle der Spontanvariation wie allgemein die Kontrolle des Zufalls. Der Arzt vergleicht mit seiner möglicherweise richtigen, möglicherweise aber auch trügerischen Vorstellung von einer Therapiewir-

kung – ein Vergleich, der der Beobachtung von Unerwartetem im Übrigen keine Chance gibt. Von dieser Trugschlussgefahr sind alle Untersuchungen ohne Kontrolle, also auch insbesondere alle Einzelfallbeurteilungen potenziell betroffen, wie überzeugend auch immer sie erscheinen. Die Medizingeschichte kennt zahlreiche Beispiele plausibler, verbreitet akzeptierter, effektvoller Behandlungsmethoden, die sich später in der empirischen Prüfung als wirkungslos erwiesen haben. Im Übrigen unterliegen auf Einzelfallbeurteilungen beruhende Therapievergleiche auch den anderen potenziellen Verfälschungsquellen, die in Abb. 7.1 aufgeführt sind. Es kann allerdings gut sein, dass dies aufgrund der fehlenden Kontrolle im konkreten Fall nicht bemerkt wird. Bei allem Respekt vor der Kunst und dem Einfühlungsvermögen des Arztes, die in der Patientenbetreuung unerlässlich sind: Allgemeine Wirksamkeitsaussagen zu therapeutischen Ansätzen lassen sich aus der unkontrollierten Einzelfallbetrachtung nicht ableiten. Die Fallstudie nimmt deshalb zu Recht in den Evidenzklassenhierarchien unterste Plätze ein.

7.3 Therapieevaluation in einer pluralen Wissenschaftswelt

Die Wissenschaftswelt ist weder im Inhalt noch in der Methodik monolithisch. Von außen betrachtet wirkt sie eher wie eine Inselwelt mit unterschiedlichen Riten und Selbstverständlichkeiten auf den einzelnen Inseln. Vom Internet wurde zeitweilig eine Intensivierung des Informationsaustausches mit beschleunigter Konsensbildung erwartet. Das Gegenteil ist der Fall: die Informationsvielfalt steigt, das Internet selbst ist im Kern nicht hierarchisch. Der Wandel wird durch das Netz noch beschleunigt. Im Heilwesen erleben wir nebeneinander eine Tendenz zur Vereinheitlichung der Qualitätsstandards, wesentlich gesteuert über die Budgetierung, und gleichzeitig das Entstehen neuer über das Netz vermittelter Gesundheits- und Informationsmärkte. Auch wissenschaftliche Ergebnisse unterliegen diesen Prozessen. Die Orientierung wird schwerer, auch durch die schiere Fülle leicht verfügbarer, aber schwer zu prüfender Information. Die Frage, wie vertrauenswürdig eine Quelle ist, gewinnt an Bedeutung. Wie kann, wie sollte die EbM auf diese Veränderungen reagieren?

In der Frühphase der EbM-Bewegung wurde häufig das Leitbild des Doktors propagiert, der mit dem internetfähigen Notebook am Krankenbett erscheint, die Diagnose, den Zustand und die Risikokonstellation des Patienten eingibt und dann die ideale Therapie nach persönlicher kritischer Würdigung der Quellen aus dem Computer zieht. Diese Vorstellung erscheint heute in Anbetracht der Informationsmenge, der Quellenvielfalt und des Fehlens eines methodischen Königswegs hoffnungslos naiv. Tatsächlich hat sich der Umgang mit der Evidenzbasierten Medizin auch anders entwickelt. Die EbM-Bewegung selbst ist heute als ein Netzwerk organisiert, in dem teilweise recht unterschiedliche Ansätze verfolgt werden, unterschiedliche Schwerpunktsetzungen und Zielvorstellungen

existieren, intensiv diskutiert und teilweise heftig gestritten wird. Die Vielfalt der Informationsansätze im Internet spiegelt die Vielfalt des EbM-Netzwerks wider. EbM entwickelt sich vom Lehrgebäude zur Austauschplattform. Dadurch wird Konsens und Critical Appraisal schwerer. Die Glaubensgewissheiten der Bewegung sind unsicherer geworden und werden selbst Gegenstand wissenschaftlicher Forschung.

Mit dieser Entwicklung erreicht die EbM-Bewegung den Reifegrad, sich dem tatsächlich vorhandenen Pluralismus in der Wissenschaftswelt zu stellen, ohne deshalb ihre Grundprinzipien aufzugeben oder beliebig zu werden. EbM steht für das Primat der empirisch begründeten Evidenz gegenüber der bloßen Behauptung. Schwierigkeiten bei der konkreten Evidenzbewertung stellen das Prinzip nicht in Frage. EbM kann mit Meinungsverschiedenheiten, Bewertungsunsicherheiten und Vorläufigkeiten leben.

Kann auch der einzelne behandelnde Arzt damit leben? Er steht unter Entscheidungszwang. Was kann ihm helfen, wenn er im Informationsdschungel und in der Deutungsvielfalt verloren zu gehen droht? Welcher Therapieempfehlung kann er trauen? Andere Bereiche unserer Gesellschaft, in denen unsichere Entscheidungen zu festen Zeitpunkten getroffen werden müssen, führen uns mögliche Modelle vor. Gerichte, Krisenstäbe, wissenschaftliche Advisory Boards haben Verfahrenskulturen entwickelt, mit denen man Verhandlungen zum Zwecke der bestmöglichen Entscheidungsfindung führen kann. In diesen Verfahren gibt es klassische Rollen – vor Gericht z.B. Anwälte, Verteidiger, Gutachter, Richter, Schöffen – die jede für sich genommen unvollkommen sind, durch die Verfahrensordnung aber zu einer gemeinsamen Problemlösung geführt werden. Die Gesamtschau ermöglicht bessere Entscheidungen. Wichtig ist dabei vor allem die Öffentlichkeit der Verhandlung. Trotz gelegentlicher Fehlurteile hat sich das

Rechtssystem bewährt und ist eine Säule unserer Staatsverfassung.

Ist die Evidenzbewertung in der Medizin möglicherweise eine vergleichbare Aufgabe? In der Konsequenz würde das bedeuten, dass wir den Methodenpluralismus nicht bekämpfen, sondern ermutigen sollten. Schräg argumentierende Staatsanwälte und Verteidiger mögen ein Ärgernis sein, sie können aber von der Spruchkammer korrigiert werden und tragen letztlich zur Rechtspflege bei. Ähnlich würden auch unkonventionelle Studienansätze zur Evidenzfindung beitragen, wenn man sie zur Beweisführung zwingt und unabhängige wissenschaftliche Spruchkammern zuließe, eine Art dritte Gewalt neben aktivem Forscher und Geldgebern.

Gibt es Gerichten vergleichbare wissenschaftliche Institutionen nicht bereits? Manche sehen die Scientific Community oder die EbM-Netzwerke und die auf Basis dieser Netzwerke erarbeiteten Reviews (Cochrane) als solche Einrichtungen an. Alternativ werden Institutionen mit eigenständigem Bewertungsauftrag (NICE, IQWiG) genannt. Politiker haben allerdings im Allgemeinen Schwierigkeiten, möglicherweise kostenträchtige Entscheidungen unabhängigen Institutionen zu überlassen, und bevorzugen daher die enge Eingrenzung des Entscheidungsspielraums durch Gesetze und Verordnungen bzw. die Überantwortung der Entscheidungen an Behörden oder an von den Politikern oder den Geldgebern dominierte Gremien. Die Evidenzfindung ist damit stets in Gefahr, von Interessengruppen instrumentalisiert zu werden. Vor diesem Hintergrund erscheint die Übertragung der Idee der Gewaltenteilung auf die Evidenzfindung in der Tat als einziger Ausweg. Eine mögliche Organisationsform wäre ein eigenständiges wissenschaftliches Gericht mit hauptberuflichen Richtern, die für die Evidenzbewertung öffentliche Verhandlungen mit Experten durchführen und abschließende Fachurteile **auf Zeit** sprechen. Die Umsetzung der Schiedssprüche könnte

dann den bestehenden Gremien (in Deutschland: der Selbstverwaltung) überlassen bleiben, die sich im Gegenzug aus der Evidenzbewertung herauszuhalten hätten. Auf der anderen Seite sollte eine solche wissenschaftliche Spruchkammer allerdings auch keine Zuständigkeiten im Bereich der Forschungsförderung haben, damit der für die Evidenzfindung wichtige Wissenschaftspluralismus nicht eingeschränkt wird, ebenso wie die Gerichte keinen direkten Einfluss auf die Arbeit der Anwälte und Gutachter ausüben sollten.

Für den Augenblick bleibt festzuhalten, dass die Vielzahl für sich genommen unvollkommener Ansätze zur Therapieevaluation für den Prozess der Evidenzgenerierung außerordentlich hilfreich ist, weil sie gegenseitige Kontrolle und kritische Bewertung anregt, vor verfrühter Festlegung bewahrt und den Prozess der Evidenzgenerierung fördert. EbM selbst ist von Natur aus ein lebensbegleitender komplexer Prozess und kein bloßes Umsetzungsprogramm und sollte als solcher gestaltet werden.

Literatur

- [1] Sox H, Screening mammography for younger women: back to basics. *Ann Intern Med.* (2002), 137, 361–362
- [2] Goodman SN, The mammography dilemma: a crisis for evidence-based medicine? *Ann Intern Med.* (2002), 137, 363–365
- [3] Evidence-Based Medicine Working Group, Evidence-based medicine. A new approach to teaching the practice of medicine. *Evidence-Based Medicine Working Group. JAMA.* (1992), 268, 2420–2425
- [4] Simpson EH, The interpretation of interaction in contingency tables. *J R Statist Soc.* (1951), Ser. B, 238–241
- [5] Rosenbaum PR (2002) *Observational studies.* Springer, Berlin
- [6] Stengel D, Porzolt F, Evidence based medicine: Randomisierung nicht zwingend erforderlich. *Dtsch Arztebl.* (2003), 100, A–2145
- [7] Kienle GS et al., Evidenzbasierte Medizin: Konkurs der ärztlichen Urteilskraft? *Dtsch Arztebl.* (2003), 100, A–2142–6

II Die Technik und das Handwerkzeug

8	Vom Problem zur Frage	89
	<i>Regina Kunz</i>	
9	Kritische Bewertung von Studien zu Prävalenz und Symptomen	93
	<i>Andreas C. Sönnichsen</i>	
10a	Kritische Bewertung von Studien zur Ätiologie	101
	<i>Torsten Schäfer</i>	
10b	Genetische Studien	115
	<i>Inke R. König, Andreas Ziegler</i>	
11	Kritische Bewertung von Studien zu diagnostischen Tests	121
	<i>Heiner C. Bucher</i>	
12	Kritische Bewertung von Studien zu therapeutischen Interventionen	133
	<i>Hilke Bertelsmann, Guido Lerzynski, Regina Kunz</i>	
13	Kritische Bewertung von systematischen Reviews und Meta-Analysen	149
	<i>Heiner C. Bucher</i>	
14	Kritische Bewertung von qualitativen Studien	159
	<i>Thorsten Meyer</i>	
15	Kritische Bewertung von Leitlinien	177
	<i>Henning Thole, Frank Thalau, Günter Ollenschläger, Ina Kopp, Monika Lelgemann</i>	
16	Kritische Bewertung von Health Technology Assessment-Berichten	191
	<i>Matthias Perleth, Dagmar Lühmann</i>	
17	Kritische Bewertung von gesundheitsökonomischen Studien	203
	<i>Reiner Leidl</i>	
18	Kritische Bewertung von Gesundheitsinformationen für medizinische Laien	217
	<i>Sylvia Säger, Marie-Luise Dierks</i>	
19	Von der Evidenz zur Empfehlung	231
	<i>Regina Kunz, Monika Lelgemann, Gordon Guyatt, Gerd Antes, Yngve Falck-Ytter, Holger Schünemann</i>	
20	Wo erhalte ich Antwort auf meine Fragen?	249
	<i>Monika Lelgemann, Norbert Donner-Banzhoff</i>	

8 Vom Problem zur Frage

Regina Kunz

- ▲ Soll man jeden Patienten mit Diabetes zum Diabetologen überweisen?
- ▲ Soll man bei V.a. Nierenarterienstenose eine digitale Subtraktionsangiographie oder eine intraarterielle Angiographie durchführen?
- ▲ Soll man bei tiefer Beinvenenthrombose niedermolekulares subkutanes oder unfraktioniertes intravenöses Heparin verwenden?

Im Klinikalltag und im Umgang mit Patienten werden wir täglich mit einer Vielzahl von Fragen und Problemen konfrontiert, und wir haben wirksame Mechanismen entwickelt, damit umzugehen und zu handeln, ohne jede Frage zu beantworten, ohne jedes Problem zu lösen. Häufig nehmen wir unsere Unwissenheit gar nicht mehr wahr. Dabei hat man eine ganz gute Vorstellung davon, wieviele patientenzentrierte Fragen in unserem Berufsalltag täglich auftauchen, deren

Beantwortung für das weitere Patienten-Management von Bedeutung wäre. So tritt in der Allgemeinarztpraxis bei durchschnittlich 15 Konsultationen eine relevante Frage auf [1], in einer allgemeinen Facharztpraxis ca. 1 Frage bei jeder 2. Konsultation [2], und bei einer Lehrvisite (Teaching ward round) für das Stationsteam werden bis zu 5 Fragen pro Patient aufgeworfen [3]. Zum Teil hängt dieses Verhalten mit unserer Ausbildung zusammen, bei der die Aussage „Ich weiß nicht“ sicher nicht zum erwünschten Repertoire gehört [4], noch wurden uns explizite Strategien vermittelt, um Fragen und Probleme zu identifizieren, zu priorisieren und effektive Lösungsstrategien zu finden. Probleme zu identifizieren und Fragen zu stellen ist ein essenzieller Bestandteil von EbM, sozusagen der Start von EbM. Ohne klar strukturierte Frage lassen sich die weiteren Schritte von EbM nicht zufriedenstellend durchführen: zur Ausgangsfrage schnell und zuverlässig

Tab. 8.1: Grundgerüst für eine gut konstruierte Frage (nach [10])

Patient/Problem	Intervention	Vergleich	Endpunkt
Mit welchen Charakteristika würde man eine Gruppe von Patienten beschreiben, die meinem Patienten ähnelt?	Welche Hauptintervention ziehe ich in Betracht?	Welche Alternativmaßnahme ziehe ich in Betracht?	Welche Ziele möchte ich mit der Maßnahme erreichen?
Wichtig ist, zwischen Genauigkeit und Kürze abzuwägen.	Präzise Angabe	Präzise Angabe	Begrenzung auf 1 oder 2 Ziele/Frage
z.B. Erziele ich bei Patienten mit einer nachgewiesenen erstmaligen proximalen tiefen Beinvenenthrombose ohne Hinweis für eine Lungenembolie durch ambulante Einleitung der Antikoagulation mit niedermolekularem Heparin verglichen mit stationärer i.v. Heparin die gleiche oder eine niedrigere Rate an Lungenembolien und/oder Mortalität?

die relevante Literatur aufzustoßern, unsere Fertigkeiten in der kritischen Bewertung auf die gefundenen Studien anzuwenden und die wissenschaftliche Zuverlässigkeit und klinische Bedeutung der Evidenz zu ermitteln, um schließlich unter Berücksichtigung unserer klinischen Erfahrung und der Wünsche und Vorstellungen des Patienten zu einer Entscheidung zu kommen. Insbesondere der Erfolg des zweiten Schritts, die Literatursu-

che, und der damit verbundene Zeitaufwand hängen neben den technischen Fähigkeiten sehr stark von einer spezifisch formulierten Frage ab. Und der benötigte Zeitaufwand wiederum bestimmt, wie häufig wir dieses Instrumentarium in der Praxis wirklich verwenden [5].

Wodurch zeichnet sich eine gut konstruierte Frage aus? Das Gerüst enthält 4 Bestandteile:

Tab. 8.2: Beispiele für die Formulierung von gut konstruierten, spezifischen Fragen für unterschiedliche Aspekte in der Versorgung von Patienten mit Thrombembolie

Klinische Aufgabe	Patient oder Problem (P)	Intervention (I), Exposition, Beobachtung	Vergleich, falls möglich (Comparison = C)	Endpunkt (Outcome = O)
Klinische Untersuchung	Beinvenenthrombose	Einseitig schmerzhaft geschwollenes Bein	Phlebographie	Genauigkeit
Aetiologie	Thrombophiles Risiko/Zustand	Pille	Keine Pille	Lungenembolie
Erstellen der Differenzialdiagnosen	Geschwollenes, gerötetes Bein	Tiefe Venenthrombose	Baker's Zyste, Erysipel, Lymphstauung	Wahrscheinlichkeit für verschiedene Ursachen
Auswahl der diagnostischen Tests	Tiefe Venenthrombose	Kompressionsultraschall	Phlebographie	Genauigkeit
Prognostische Vorhersage	APC-Resistenz oder positiver Leidenfaktor		Keine APC-Resistenz	Erstmanifestation einer Thrombose
Auswahl der Therapie	Proximale Tiefe Venenthrombose	Stationäre Behandlung mit unfraktioniertem Heparin	Ambulante Behandlung mit niedermolekularem Heparin	Lungenembolie
Auswahl präventiver Maßnahmen	Beinvenenthrombose	6 Monate Antikoagulation	Zeitlich unbegrenzte Antikoagulation	Rezidiv
Beurteilung der Nebenwirkungen	Beinvenenthrombose	Antikoagulation	Keine Antikoagulation	Blutungsereignisse
Qualitätsverbesserung	Verfügbarkeit der D-Dimer-Testung bei V.a. Thrombembolie	Bedside-Test rund um die Uhr	nur während Routinelaborzeiten	Ressourcenverbrauch
Verbesserung des eigenen Lernverhaltens	Niedermolekulares Heparin	Gang in die Bibliothek	Internet oder CD-Rom	Zeitaufwand und Verständnis

Tab. 8.3: Orientierungsraster für die Eingrenzung des Patientenproblems und zur Konstruktion einer PICO-Frage

Handelt es sich bei meinem Problem um ...
<ul style="list-style-type: none"> • eine klinische Information – wie man klinische Befunde aus Anamnese und klinischer Untersuchung gezielt erhebt und sie richtig interpretiert? • Diagnostik – wie man diagnostische Tests auswählt und interpretiert? • die Prognose – wie man den weiteren Verlauf bei dem jeweiligen Patienten richtig einschätzt ? • therapeutische Maßnahmen – wie man Therapien und Interventionen auswählt, sodass man sicher sein kann, dass sie mehr nützen als schaden? • Prävention – wann und wie man Screening-Tests durchführt und Krankheitsrisiken vermindert? • Schulung und Fortbildung – wie man sowohl dem Patienten und seiner Familie als auch sich selbst das nötige Wissen, Kenntnisse und Fertigkeiten vermittelt.

- ▲ Eine Beschreibung des Patienten/des Problems,
- ▲ die Intervention (wobei der Begriff sehr weit gefasst auch eine Ursache, einen prognostischen Faktor oder eine Behandlung bedeuten kann),
- ▲ die Vergleichsmaßnahme und
- ▲ die Endpunkte, an denen man den Erfolg messen möchte.

Diese Struktur hat sich in der EbM-Terminologie unter dem Begriff „PICO“-Frage fest etabliert.

Die gut konstruierte Frage soll zum einen das individuelle Patientenproblem (das den Prozess schließlich ausgelöst hat) widerspiegeln, zum anderen sollen sich daraus klare Suchstrategien ergeben, die zu entsprechend relevanten und präzisen Antworten führen.

Mit diesem Gerüst lassen sich die meisten Probleme und Fragen, die uns im klinischen Alltag begegnen, in Schlüsselbegriffe fassen (vgl. Tab. 8.2), mit denen sich die anschließende Literatursuche schnell und spezifisch durchführen lässt – unabhängig davon, welches Suchsystem und welche Datenbank man verwendet. Allerdings benötigt das Formulieren von Fragen, wie alle klinischen Fertigkeiten, etwas Übung und regelmäßige Anwendung. Gerade EbM-Anfänger finden es zu Beginn schwierig, sich spezifisch festzulegen und die eigene Frage von anderen ähnlichen Fragen, die aber einen

anderen Fokus haben, abzugrenzen. Um sich über die Zielrichtung der eigenen Frage klar zu werden, hat sich das Raster von EbM-Kollegen (Mark Wilson, Scott Richardson [6]) bewährt (s. Tab. 8.3).

Hat man erst einmal einen Blick für Fragen und ungelöste Probleme entwickelt und sich auf die damit verbundenen Unsicherheiten in der eigenen Praxis eingelassen, so fühlt man sich angesichts des auftauchenden Fragenbergs schnell überwältigt. Im Folgenden sollen einige Strategien aufgezeigt werden, wie sich identifizierte Probleme auf machbare Weise lösen lassen:

EbM kann natürlich allein praktiziert werden, aber es macht immer mehr Spaß und fördert die Kommunikation und den Austausch, mit einem oder mehreren Kollegen EbM-Lösungen zu entwickeln. Dafür eignen sich Stationsvisiten und Abteilungsfortbildungen, Qualitätszirkel, Gemeinschaftspraxen oder Praxisgemeinschaften sowie Praxisnetzwerke.

Eine Möglichkeit ist, über einen Zeitraum (z.B. von 2 Wochen) alle Fragen, die während der Arbeit auftauchen, zu sammeln und dann eine Prioritätenliste aufzustellen, in welcher Reihenfolge man die Fragen abarbeiten möchte. Dies kann unter unterschiedlichen Gesichtspunkten geschehen: nach der Dringlichkeit des Problems (z.B. Antwort sofort notwendig für die Behandlung des Patienten), seiner Häufigkeit, dem persönli-

chen Interesse, der verfügbaren Zeit zur Lösungsfindung, der klinischen Bedeutung ... In der Klinik oder in der ambulanten Weiterbildung kann man mit den Ärzten in Weiterbildung Übungen vereinbaren [7; 8], bei denen sie in einem vorgegebenen Zeitraum ein definiertes Patientenproblem mithilfe der Literatur lösen und das Resultat dem Stations-team zurückmelden sollen. Die gefundenen Lösungen lassen sich mithilfe eines einfach zu handhabenden Computerprogramms in einseitige so genannte CATs (Critically appraised topics) verpacken [9] und ausdrucken. Damit kann man eine für die Probleme vor Ort maßgeschneiderte Studiensammlung anlegen, auf die bei erneutem Auftauchen der Frage innerhalb von Sekunden zurückgegriffen werden kann [6].

Steht ein Management-Problem allgemeiner Art im Vordergrund, lohnt sich eine Suche nach evidenzbasierten Leitlinien (s. Kap. 35), die entweder den Gegebenheiten vor Ort entsprechen oder angepasst werden.

Mit etwas Motivation und Organisations-talent lässt sich so in den täglichen Berufs- und Praxisalltag eine hochwirksame Strategie einbauen, mit der man sich bedarfsorientiert auf dem Laufenden hält, die zur Kommunikation mit Kollegen einlädt und die nicht zuletzt die Neugierde und die Freude an der Medizin aufrecht erhält.

Literatur

- [1] Ely JW, Burch RJ, Vinson DC, The information needs of family physicians: case-specific clinical questions. *J Fam Pract.* (1992), 35, 265–269
- [2] Covell DG, Uman GC, Manning PR, Information needs in office practice. Are they being met? *Ann Intern Med.* (1985), 103, 96–99
- [3] Osheroff JA et al., Physicians' information needs: analysis of questions posed during clinical teaching. *Ann Intern Med.* (1991), 114, 576–581
- [4] „I don't know“: the three most important words in education. *BMJ.* (1999), 318, A
- [5] Sackett DL, Straus SE, Finding and applying evidence during clinical rounds: the „evidence cart“. *JAMA.* (1998), 280, 1336–1338
- [6] (1997) Workshop: How to teach Evidence-based Medicine. McMaster Univ, Hamilton
- [7] Sackett DL, Kunz R (1999) Evidenzbasierte Medizin. EbM-Umsetzung und -Vermittlung (Deutsche Ausgabe). Zuckschwerdt, Bern u.a.
- [8] Centre for Evidence-Based Medicine. Educational Prescriptions. <http://www.cebm.net/eduscrip.asp> (09.0107)
- [9] Centre for Evidence-Based Medicine. CAT-maker and CATbank. <http://www.cebm.net/cats.asp> (09.0107)
- [10] Centre for Evidence-Based Medicine. Focusing Clinical Questions. http://www.cebm.net/focus_quest.asp (09.01.07)

9 Kritische Bewertung von Studien zu Prävalenz und Symptomen

Andreas C. Sönnichsen

Der typische Anlass für die Inanspruchnahme eines Arztes ist das Symptom. Der Patient bemerkt an sich selbst eine Veränderung, die ihm Sorge bereitet. Hierbei ist zunächst unbedeutend, ob die Veränderung durch Beschwerden, also beispielsweise Schmerzen oder sonstige unangenehme Wahrnehmungen gekennzeichnet ist, oder ob lediglich eine Abweichung vom sonst Gewohnten vorliegt. Die Veränderungen können physischer oder psychischer Natur sein, akut auftreten oder sich langsam entwickeln, bis für den Patienten ein Schwellenwert erreicht ist, der ihn zum Aufsuchen eines Arztes veranlasst. An den Arzt wird dann die Forderung herangetragen, die Ursachen des Symptoms zu ergründen, es unter Zuhilfenahme diagnostischer Maßnahmen in den Kontext einer eventuell vorhandenen Krankheit einzuordnen, eine Diagnose zu stellen, erforderlichenfalls eine Behandlung einzuleiten und eine Aussage zur Prognose zu treffen. Die Bewertung von Symptomen mit den daraus abgeleiteten Konsequenzen gehört somit zur zentralen Aufgabe des Arztes und fordert daher eine evidenzbasierte und wissenschaftlich fundierte Betrachtung. Das Symptom kann hierbei im weitesten Sinne als diagnos-

tischer Test betrachtet werden, der die Wahrscheinlichkeit für das Vorliegen einer dem Symptom zugrunde liegenden Erkrankung angibt (positiver prädiktiver Wert). Diese Wahrscheinlichkeit ist gleichbedeutend mit der Prätest-Wahrscheinlichkeit für nachgeschaltete weitere Diagnostik.

Die Fragen, die in Studien zur Evaluation von Symptomen bearbeitet werden müssen, sind in Tabelle 9.1 dargestellt.

9.1 Klinisches Szenario

Herr Dr. Ungewiss ist ein engagierter Allgemeinarzt in einem Stadtviertel mit hohem Arbeiter- und Ausländeranteil. Fast täglich kommt ein Patient wegen Kreuzschmerzen in seine Sprechstunde. Einige Patienten überweist der Kollege zum Röntgen der Lendenwirbelsäule, andere nicht. Er lässt sich dabei vor allem von seinem „Bauchgefühl“ und der vom Patienten angegebenen Intensität des Schmerzes leiten. In den meisten Fällen werden vom Radiologen „degenerative Veränderungen“ konstatiert, ohne dass sich hieraus eine Konsequenz für das weitere Vorgehen ergibt. Je nach Schmerzintensität werden die Patienten unterschiedlich lange krank geschrieben, erhalten nicht-

Tab 9.1: Schritte zur evidenzbasierten Bewertung von Symptomen

1	In welcher Zielpopulation wird das Symptom untersucht?	Setting
2	Mit welcher Häufigkeit kommt das Symptom in der Zielpopulation vor?	Prävalenz des Symptoms
3	Welche Störungen liegen dem Symptom mit welcher Häufigkeit zugrunde?	Prävalenz der möglichen Erkrankungen unter den Symptompositiven (Differenzialdiagnose)
4	Welcher Verlauf ist bei Vorliegen des Symptoms zu erwarten?	Prognose

steroidale Antirheumatika (NSAR) und Physiotherapie. Herr Dr. Ungewiss ist zunehmend unzufrieden mit dem eigenen unsystematischen Vorgehen und möchte als ersten Schritt auf dem Weg zu einer strukturierten Versorgung von Patienten mit Kreuzschmerzen in Erfahrung bringen, welche Differenzialdiagnosen mit welcher Häufigkeit bei Patienten mit Kreuzschmerzen in der allgemeinärztlichen Praxis vorkommen.

9.2 Die Evidenz suchen

Herr Dr. Ungewiss recherchiert zunächst in einem allgemeinmedizinischen Lehrbuch und findet für Kreuzschmerzen folgende Differenzialdiagnosen:

- ▲ Unkomplizierte, funktionelle Schmerzen (idiopathische Kreuzschmerzen),
- ▲ Bandscheibenschäden (Protrusion, Prolaps),
- ▲ Spinalkanalstenose,
- ▲ Osteoporose bzw. osteoporotische Fraktur,
- ▲ Trauma,
- ▲ Entzündliche Erkrankungen (Spondylodisitis, Osteomyelitis),
- ▲ Rheumatische Erkrankungen (Spondylitis ankylosans, Sakroiliitis),
- ▲ Tumorerkrankungen (Skelettmetastasen, retroperitoneale Tumoren),
- ▲ Urologische Erkrankungen (Pyelonephritis, Urolithiasis),
- ▲ Gynäkologische Erkrankungen (Endometriose, Lageanomalien des Uterus),
- ▲ Interne Erkrankungen (Pankreatitis, Divertikulitis, Cholezystitis, Cholelithiasis).

Die Fülle der möglichen Differenzialdiagnosen macht deutlich, dass es unmöglich ist, in allen Fällen durch Labordiagnostik, bildgebende Verfahren und sonstige, vielleicht sogar invasive diagnostische Maßnahmen sämtliche in Frage kommenden Diagnosen auszuschließen.

Aber: Wann ist eine weitere Abklärung indiziert und wann nicht? Um auf diese Fra-

ge eine Antwort zu finden, ist es notwendig, das Symptom oder den Symptomenkomplex „Kreuzschmerzen“ in differenzialdiagnostischen Symptomstudien zu untersuchen.

Um in der medizinischen Literatur nach einer solchen Studie zu suchen, ist es hilfreich, die gewünschte Studienfragestellung selbst zu formulieren. Herr Kollege Ungewiss formuliert also folgende Frage: Welche Differenzialdiagnosen stellen sich nach längerer Beobachtung (z.B. innerhalb eines Jahres) bei Patienten heraus, die wegen akuter Kreuzschmerzen ihren Hausarzt aufsuchen?

Leider ist die Datenlage hierzu insgesamt dünn, vor allem, wenn man nach Studien sucht, die sich mit Häufigkeit, Differenzialdiagnose und Prognose von Symptomen im primärärztlichen Bereich beschäftigen. Aus diesem Grunde gestaltet sich auch die weitere Literatursuche des Kollegen Ungewiss schwierig. Die MEDLINE-Suche mit den Termini „**low back pain AND (primary care OR general practice) AND (differential diagnosis OR outcome) AND symptom* AND predictive**“ ergibt nur sechs Treffer. Eine der gefundenen Studien untersucht unsere Fragestellung, aber nur an 103 Patienten mit ausschließlich unkomplizierten Kreuzschmerzen [1]. Die Patienten, die uns interessieren, bei denen sich nämlich hinter den Kreuzschmerzen eine ernsthafte Erkrankung verbirgt, waren in dieser Studie zufällig aufgrund der niedrigen Fallzahl nicht dabei.

Wir fassen die Suche also weiter: Die Termini „**low back pain AND (primary care OR general practice) AND (differential diagnosis OR prevalence OR epidemiology)**“ bringen zwar 245 Treffer, aber beim Durchsuchen stellt sich heraus, dass eigentlich keine Studie dabei ist, die wirklich genau das untersucht, was wir wissen wollen.

Wir behelfen uns also mit Studien, die zumindest teilweise die gewünschten Informationen beinhalten. So stoßen wir beispielsweise auf die Untersuchung von Schers et al. [2]. Ziel der Studie war eigentlich, das Management

Tab 9.2: Häufigkeit der zugrunde liegenden Differenzialdiagnosen von Patienten mit Kreuzschmerzen in der hausärztlichen Praxis

Diagnose	N	% der Pat.
Unspezifische Kreuzschmerzen	1180	71,9
Radikuläre Schmerzen durch Bandscheibenschaden oder sonstige Wurzelkompression	234	14,3
Maligne Tumorerkrankung	23	1,4
Fraktur	34	2,1
Sonstige Diagnose	169	10,3
Konsultationen gesamt	1640	100,00

Nach Schers H et al. [2]

von Kreuzschmerzen in hausärztlichen Praxen hinsichtlich der Anwendung von Leitlinien zu untersuchen. Als Nebenprodukt erfahren wir aber die dem Symptom Kreuzschmerzen zugrunde liegenden Diagnosen (s. Tab. 9.2).

9.3 Die Evidenz bewerten – kritische Evaluation einer Studie zur Prävalenz, Differenzialdiagnose und Prognose eines Symptoms

Betrachten wir die Studie von Schers näher hinsichtlich unserer Fragestellung und versuchen wir, ihre Validität zu beurteilen. Hierzu analysieren wir die Untersuchung hinsichtlich der Qualitätskriterien für Symptom-Evaluationsstudien in der Primärmedizin, wie sie zusammenfassend von Donner-Banzhoff et al. dargestellt wurden [3].

9.3.1 Auswahl der Patienten

Die valide Erfassung der Häufigkeit bestimmter Diagnosen bei Vorliegen eines bestimmten Symptoms ist nur möglich, wenn innerhalb eines definierten Beobachtungszeitraums alle infrage kommenden Patienten mit dem zu untersuchenden Symptom in die Studie aufgenommen werden. Dies ist selbst bei bester Absicht häufig schwierig. Im hekti-

schen Praxisalltag passiert es leicht, dass harmlose, unspektakuläre Fälle übersehen werden und nicht in die Studie gelangen. Dies führt zu einer Überrepräsentation schwerer Erkrankungen in der differenzialdiagnostischen Abbildung. Wir fordern also nicht nur eine prospektive und lückenlose konsekutive Aufnahme von allen Patienten mit dem betreffenden Symptom, sondern auch ein zumindest stichprobenweise durchgeführtes externes Monitoring der teilnehmenden Praxen, um die lückenlose Rekrutierung sicherzustellen bzw. durch die Dokumentation der Anzahl der „übersehenen“ Fälle einen Quotienten aus den tatsächlich rekrutierten Patienten und denen, die hätten rekrutiert werden müssen, zu bilden.

Als weitere Forderung ist es notwendig, Patienten nur einmal in die Studie aufzunehmen. Das heißt, Patienten, die während des Rekrutierungszeitraums mehrfach mit dem gleichen (zu untersuchenden) Symptom in die Praxis kommen, gehen nur einmal in die Studie ein. Es liegt auf der Hand, dass auch hier die mehrfache Aufnahme ein und desselben Patienten zu einer Überbewertung schwerer Fälle führen würde.

Die lückenlose Rekrutierung war in der Studie von Schers zwar beabsichtigt, wurde jedoch mit großer Wahrscheinlichkeit nicht eingehalten. An der Untersuchung waren 30 Praxen beteiligt, der Rekrutierungszeitraum betrug 8 Monate, die Zahl der rekrutierten

Patienten lag bei 1 390. Dies entspricht einer Anzahl von knapp sechs Rückenschmerz-Patienten pro Praxis und Monat, was in deutlicher Diskrepanz zu Untersuchungen über die Häufigkeit von Kreuzschmerzen in der Allgemeinpraxis steht (mittlere Inzidenz von 5,5% der Patienten) [4]. Ein entsprechendes Monitoring zur Verminderung oder zumindest quantitativen Erfassung nicht rekrutierter Patienten ist nicht erfolgt. Als weiterer Kritikpunkt ist die Auswertung von Patientenkontakten statt Patienten anzumerken. Die 1 640 analysierten Patientenkontakte stammen von nur 1 396 Patienten. Die Häufigkeit der Diagnosen wird aber auf die Patientenkontakte bezogen und wir erfahren nicht, welche Patienten bzw. welche Diagnosen hier doppelt gezählt wurden.

Beide Fehler – die nicht konsekutive Rekrutierung und die Erfassung von Kontakten statt Patienten – führen zu einer Überrepräsentation der ernstesten Erkrankungen. Nur so ist die hohe Zahl an radikulären Kreuzschmerzpatienten (14%) und die relativ hohe Zahl maligner Tumoren (1,4%) zu erklären. In einer älteren Studie wird der Anteil an Patienten mit signifikanten radikulären Symptomen (z.B. durch Bandscheibenschäden) im allgemeinmedizinischen Setting mit 1,4% angegeben, Tumorerkrankungen als Ursache von Kreuzschmerzen sind in der Allgemeinpraxis mit <0,7% eine Rarität [5; 6].

9.3.2 Einschlusskriterien

Für die Evaluation eines Symptoms ist es entscheidend, welche Patienten in die Studie eingeschlossen werden. Eine Altersbegrenzung führt zu einer Veränderung des Studienergebnisses, da Krankheitswert und Zusammenhang mit bestimmten Diagnosen meist altersabhängig variieren. Auch ist es wichtig, ob nur Patienten eingeschlossen werden, die explizit wegen des untersuchten Symptoms in die Praxis kommen, oder ob

die Patienten unabhängig vom Behandlungsanlass gezielt nach dem zu untersuchenden Symptom gefragt werden, was dann im positiven Fall zur Aufnahme in die Studie führt.

Für Kreuzschmerzen besteht beispielsweise eine sehr hohe jährliche Inzidenz in der westlichen Bevölkerung (bis zu 40%/Jahr) [7]. Nur ein geringer Anteil sucht deshalb einen Arzt auf. Das bedeutet, dass das gezielte Nachfragen nach Kreuzschmerzen bei Patienten, die aus anderen Gründen in die Praxis kommen, zu einem Anstieg der Inzidenz von Kreuzschmerzen in der Allgemeinpraxis führt.

Dennoch können beide Vorgehensweisen in der Evaluation von Symptomen sinnvoll sein. Zur Erfassung epidemiologischer Merkmale eines Symptoms müssen möglichst alle Patienten, die das Symptom haben, erfasst werden. Für die Symptombeurteilung in der Praxis des Allgemeinmediziners kann es dagegen besser sein, nur die Patienten zu rekrutieren, die tatsächlich wegen des Symptoms die Praxis aufsuchen. In jedem Fall sollte eine Studie zur Symptomevaluation die Vorgehensweise angeben und in der Diskussion der Ergebnisse aufgreifen. In unserer Beispielstudie wird hierzu nichts gesagt. Weiterhin muss geklärt werden, ob nur Patienten, bei denen das Symptom zum ersten Mal auftritt, in die Studie aufgenommen werden, oder ob auch diejenigen eingeschlossen werden, bei denen es zwar zu einer neuen Episode gekommen ist, die aber das Symptom schon öfter hatten.

Ein weiterer Aspekt, der festgelegt werden muss, ist die Symptomdefinition und die unter Umständen erforderliche Definition eines Schwellenwerts. Gerade beim Symptom Schmerz gestaltet sich diese Definition schwierig. Zum einen ist das Schmerzempfinden der Patienten ausgesprochen subjektiv. Zum anderen werden als gleich intensiv empfundene Schmerzen den einen Patienten zum Arzt führen – vielleicht nur,

um eine Arbeitsunfähigkeitsbescheinigung zu bekommen –, während der andere seine Schmerzen zu Hause selbst auskuriert. Auch die – ebenfalls subjektive – Bewertung der geschilderten Symptomatik durch den Arzt kann zu einer unterschiedlichen Handhabung der Einschlusskriterien führen. Diese müssen also so exakt wie möglich definiert werden, und die an der Studie teilnehmenden Ärzte sind in der einheitlichen Handhabung zu schulen.

9.3.3 Studienpopulation

Prävalenz und Umgang mit einem Symptom in der Bevölkerung kann nach Region und sozialem Status stark variieren. So kann das Bildungsniveau sowohl die Inzidenz als auch das Outcome von Kreuzschmerzen beeinflussen [8]. Das spezifische Setting einer Studie zur Evaluation von Symptomen kann also das Ergebnis nachhaltig verändern. Dem kann man durch Beteiligung multipler Praxen in unterschiedlichen Settings begegnen. Auch der Einflussfaktor „Arzt“, der sich auf das zur Disposition stehende Patientenkollektiv und auf das Rekrutierungsverhalten auswirkt, wird durch Multizentrität minimiert.

9.3.4 Denominator

Der Denominator oder Nenner, auf den sich eine Studie zur Symptomevaluation bezieht, ist für das Studienergebnis von entscheidender Bedeutung. In manchen Ländern, wie z.B. in Großbritannien oder Dänemark, müssen sich alle Personen – auch die Gesunden – für einen bestimmten Hausarzt registrieren lassen. Diese Patientenlisten des an einer Studie teilnehmenden Arztes können als Nenner verwendet werden, um die Zahl der symptomatischen Patienten auf die Gesamtzahl der von diesem Arzt betreuten Personen

zu beziehen. In Deutschland oder Österreich ist dies nicht möglich. Gesunde tauchen in der Statistik des Hausarztes nicht auf. Es wurde verschiedentlich vorgeschlagen, in diesen Ländern die Jahreskontaktgruppe als Denominator zu verwenden [9]. In Studien zur Prävalenz von Symptomen muss daher angegeben werden, auf welchen Denominator sich die Studienergebnisse beziehen.

9.3.5 Asymptomatische Vergleichsgruppe

Aus mehreren Gründen kann es sinnvoll sein, gleichzeitig mit den symptomatischen auch asymptomatische Patienten zu untersuchen. Wenn wir das Vorliegen eines Symptoms als diagnostischen Test begreifen, dessen Sensitivität und Spezifität für das Vorliegen einer bestimmten Erkrankung erfasst werden soll, so ist es erforderlich, auch alle Symptom-Negativen in die Studie aufzunehmen, die im Rekrutierungszeitraum in den an der Studie teilnehmenden Allgemeinpraxen vorstellig werden (das entspricht dem asymptomatischen Anteil des Denominators). Um beim Beispiel Kreuzschmerzen zu bleiben, kann es von Interesse sein, die Häufigkeit von asymptomatischen Patienten mit nachgewiesenem Bandscheibenschaden zur Häufigkeit von Bandscheibenschäden bei Patienten mit Kreuzschmerzen in Beziehung zu setzen. So wird in einer Arbeit berichtet, dass bei mehr als 20% asymptomatischer Patienten in einer Magnetresonanztomographie ein relevanter Diskusprolaps oder eine Spinalkanalstenose nachgewiesen werden konnte [10]. Analog wird man bei Kreuzschmerzpatienten bei routinemäßiger Durchführung einer Magnetresonanztomographie zahlreiche positive Befunde erheben, die wahrscheinlich mit den Kreuzschmerzen nichts zu tun haben. Die Evaluation eines Symptoms nur in einer Gruppe symptomatischer Patienten kann also zu einer Überre-

präsentation vermeintlicher Diagnosen führen. Schlimmstenfalls werden in unserem Beispiel zahlreiche Patienten der operativen Therapie eines asymptomatischen Bandscheibenvorfalls zugeführt, während die damit nicht assoziierten Kreuzschmerzen fortbestehen.

Der andere Grund für die gleichzeitige Untersuchung asymptomatischer Patienten ist, die Häufigkeit assoziierter Symptome, Erkrankungen und möglicher kausaler Faktoren zu erfassen. Wenn man beispielsweise den kausalen oder prognostischen Faktor „Depression“ hinsichtlich seiner Häufigkeit bei Kreuzschmerzpatienten untersucht, so ist diese Aussage nur verwertbar, wenn man gleichzeitig auch die Häufigkeit von Depressionen bei Patienten ohne Kreuzschmerzen erfasst.

Methodisch kann die asymptomatische Gruppe auf zweierlei Weise gebildet werden. Sensitivität und Spezifität des untersuchten Symptoms lassen sich nur durch Untersuchung der vollständigen Population (z.B. aller im Studienzeitraum in den teilnehmenden Praxen behandelten Patienten) gewinnen. Für die Berechnung der prädiktiven Werte und der prozentualen Häufigkeit von Erkrankungen genügt ein Fall-Kontroll-Ansatz.

9.3.6 Einheitliches diagnostisches Vorgehen

Wie wir bereits gesehen haben, führt die Durchführung einer diagnostischen Maßnahme, z.B. Magnetresonanztomographie (MRT), zur Entdeckung versteckter Diagnosen, die möglicherweise mit dem untersuchten Symptom nichts zu tun haben. Das diagnostische Vorgehen muss daher in einer Studie zur Symptomevaluation transparent, strukturiert und einheitlich erfolgen. Die Überdetektion von Erkrankungen durch vermehrte Diagnostik in der Gruppe sympto-

matischer Patienten wird als „Unmasking“- oder „Detection signal bias“ bezeichnet [11].

9.3.7 Eindeutigkeit diagnostischer und prognostischer Definitionen

Für die Verwertbarkeit einer Studie zur Evaluation von Symptomen ist es notwendig, dass die in der Studie erfassten Diagnosen und Prognosen eindeutig definiert werden. So wird beispielsweise die Prognose von Patienten mit Kreuzschmerzen in einer Studie in „poor“, „fair“ und „good outcome“ eingeteilt, wobei die Definitionen vage bleiben [12]. In einer anderen Arbeit wird die Prognose hinsichtlich der Schmerzen mit einer visuellen Analogskala und hinsichtlich der Funktionalität mit dem Hanover back pain daily activity schedule untersucht, wodurch valide und reproduzierbare prognostische Aussagen möglich werden [13]. Eine generelle Standardisierung zur Beurteilung des Outcomes von Kreuzschmerzpatienten wurde empfohlen und sollte für alle Evaluationsstudien von Symptomen gefordert werden [14].

9.3.8 Nachbeobachtung zur Sicherung von Diagnose und Prognose

Wenn man ein Symptom hinsichtlich der Stellung einer definitiven Diagnose evaluieren möchte, ist eine hinreichend lange Nachbeobachtungszeit der Patienten erforderlich, da meist nicht alle Patienten einer ausgedehnten diagnostischen Abklärung zugeführt werden können. Zum einen würde dies, wie bereits dargestellt, zur „Produktion“ von Diagnosen führen, die mit dem Symptom nichts zu tun haben, zum anderen sind eventuell erforderliche (invasive) diagnostische Maßnahmen nicht allen Patienten zumutbar und außerdem aus wirtschaftlichen Gründen nicht möglich. Häufig wird sich also eine Diagnose erst durch Beobachtung

des klinischen Verlaufs herausstellen. In der Studie von Schers wird der Zeitraum dieses Follow-ups nicht angegeben, was die Sicherheit hinsichtlich der angegebenen Diagnosen in Frage stellt [2]. Hinsichtlich der Diagnosestellung wird eine Nachbeobachtung von einem halben bis einem Jahr als ausreichend angesehen.

9.3.9 Nachbeobachtung zur Evaluation der Prognose

Für eine Aussage zur Prognose sind meist längere Beobachtungszeiträume erforderlich. Die symptomatischen Patienten werden dann im Sinne einer Kohortenstudie weiter verfolgt, um das Langzeitergebnis des untersuchten Symptoms zu erfassen.

Wichtig für die prognostische Nachbeobachtung ist auch, dass sich die in die Studie eingeschlossenen Patienten in einem vergleichbaren Krankheits- bzw. Symptomstadium befanden. So macht es beispielsweise wenig Sinn, Patienten mit neu aufgetretenen Kreuzschmerzen und Patienten mit bereits chronifizierten Kreuzschmerzen zusammen zu untersuchen.

9.3.10 Fallzahlberechnung

Die Studie von Schers macht deutlich, dass die erforderliche Fallzahl zur Evaluation von Symptomen im primärmedizinischen Bereich häufig ein Problem darstellt, weil die für die Fragestellung relevanten ernsten Erkrankungen nur sehr selten vorkommen. So liegt das 95%-Konfidenzintervall für die Häufigkeit von malignen Tumoren als Ursache der Kreuzschmerzen in dieser Studie trotz der hohen Fallzahl von 1640 mit $1,4 \pm 0,6\%$ (0,8–2,0%) nur gerade noch im akzeptablen Bereich [2]. Die erforderliche Fallzahl für die Evaluation eines Symptoms hinsichtlich des Vorliegens einer bestimmten Erkrank-

ung hängt von der erwarteten Prävalenz der Erkrankung unter den Symptompositiven und von der angestrebten Genauigkeit (ausgedrückt als Konfidenzintervall) ab.

9.3.11 Statistische Analyse

Wie bereits unter 9.3.10 dargestellt, ist die bloße Prozentangabe über die Häufigkeit von Diagnosen oder bezüglich eines bestimmten prognostischen Ausgangs nicht hinreichend, um die Relevanz von Symptomen zu beurteilen. Hier sollten immer die 95%-Konfidenzintervalle angegeben werden, um die Spannweite der möglichen Aussage und damit ihre Genauigkeit abzubilden.

9.4 Auflösung des klinischen Szenarios

In Tabelle 9.3 sind alle Qualitätskriterien für Studien zur Evaluation von Symptomen zusammenfassend dargestellt. Wie wir unschwer erkennen können, erfüllt die Beispielstudie zum Thema Kreuzschmerzen kaum eines dieser Qualitätskriterien. Aus der untersuchten Studie, aber auch aus den anderen zitierten Studien über Prävalenz, Differenzialdiagnose und Prognose von Kreuzschmerzen können nur wenige sichere Rückschlüsse gezogen werden, die die Frage des Kollegen Ungewiss mit hinreichender Genauigkeit beantworten. Lediglich die Zusammenschau der gesamten verfügbaren Literatur lässt erahnen, dass schwere Erkrankungen als Ursache von Kreuzschmerzen so selten sind, dass ohne zusätzliche, auf eine schwere Erkrankung hinweisende Symptome oder anamnestische Angaben weitere (z.B. bildgebende) Diagnostik nicht gerechtfertigt ist. Die Studienlage macht außerdem deutlich, dass bezüglich der Evaluation von Symptomen in der Primärmedizin erheblicher Forschungsbedarf besteht.

Tab. 9.3: Qualitätskriterien für Studien zur Evaluation von Symptomen

Qualitätskriterium	Fragen
Auswahl der Patienten	Wurden alle konsekutiven Patienten aufgenommen? Wurde die konsekutive Aufnahme durch Monitoring kontrolliert?
Einschlusskriterien	Wurden die Einschlusskriterien definiert? Wurden nur Patienten aufgenommen, die wegen des Symptoms in die Praxis kamen oder wurde gezielt nach dem Symptom gefragt?
Studienpopulation	Wurde die Studie multizentrisch durchgeführt? Ist die Studienpopulation repräsentativ für die Gesamtbevölkerung?
Denominator	Welcher Nenner wurde als Bezugsgröße für die Symptompositiven gewählt?
Asymptomatische Vergleichsgruppe	Wurde eine asymptomatische Kontrollgruppe untersucht? Wenn ja, als Kohorten- oder als Fall-Kontroll-Studie?
Einheitliches diagnostisches Vorgehen	Wurden alle Patienten einem standardisierten diagnostischen Vorgehen unterworfen?
Eindeutigkeit diagnostischer und prognostischer Definitionen	Wurden Diagnosen und Prognosen eindeutig und reproduzierbar definiert?
Nachbeobachtung zur Sicherung der Diagnose	Wurden die Patienten ausreichend lange nachbeobachtet, um die angegebenen Diagnosen mit Sicherheit zu stellen?
Nachbeobachtung zur Evaluation der Prognose	Wurden die Patienten ausreichend lange nachbeobachtet, um die angegebenen Prognosen machen zu können?
Fallzahlberechnung	Ist die Fallzahl ausreichend hoch, um bei Angabe des 95%-Konfidenzintervalls sinnvolle Aussagen machen zu können?
Statistische Analyse	Wurden die 95%-Konfidenzintervalle durchgehend angegeben?

Literatur

- [1] Coste J et al., Clinical course and prognostic factors in acute low back pain: an inception cohort study in primary care practice. *BMJ*. (1994), 308, 577–580
- [2] Schers H et al., Low back pain in general practice: reported management and reasons for not adhering to the guidelines in The Netherlands. *Br J Gen Pract*. (2000), 50, 640–644
- [3] Donner-Banzhoff N, Kunz R, Rosser W, Studies of symptoms in primary care. *Fam Pract*. (2001), 18, 33–38
- [4] Becker A et al., Kreuzschmerzpatienten in der hausärztlichen Praxis: Beschwerden, Behandlungserwartungen und Versorgungsdaten. *Z Allg Med*. (2003), 79, 126–131
- [5] Barton JE et al., Low back pain in the primary care setting. *J Fam Pract*. (1976), 3, 363–366
- [6] Deyo RA, Diehl AK, Cancer as a cause of back pain: frequency, clinical presentation, and diagnostic strategies. *J Gen Intern Med*. (1988), 3, 230–238
- [7] Walsh K, Cruddas M, Coggon D, Low back pain in eight areas of Britain. *J Epidemiol Community Health*. (1992), 46, 227–230
- [8] von Korff M et al., Back pain in primary care. Outcomes at 1 year. *Spine*. (1993), 18, 855–862
- [9] Szecsenyi J et al., Eine Methode zur Bestimmung des Denominators in Allgemeinpraxen – Ergebnisse einer Pilotstudie. *Gesundheitswesen*. (1993), 55, 32–36
- [10] Boden SD et al., Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am*. (1990), 72, 403–408
- [11] Sackett DL, Bias in analytic research. *J Chronic Dis*. (1979), 32, 51–63
- [12] Schiottz-Christensen B et al., Long-term prognosis of acute low back pain in patients seen in general practice: a 1-year prospective follow-up study. *Fam Pract*. (1999), 16, 223–232
- [13] Croft PR et al., Outcome of low back pain in general practice: a prospective study. *BMJ*. (1998), 316, 1356–1359
- [14] Deyo RA et al., Outcome measures for low back pain research. A proposal for standardized use. *Spine*. (1998), 23, 2003–2013

10a Kritische Bewertung von Studien zur Ätiologie

Torsten Schäfer

Die Frage nach der Ätiologie von Erkrankungen soll anhand des Beispiels der Hormonerersatz-Therapie und Brustkrebs besprochen werden. Dafür werden eine Fall-Kontroll- und eine Kohortenstudie besprochen und diese Studientypen auch aus genetischer Sicht beleuchtet.

10a.1 Klinisches Szenario

Sie sind Gynäkologe/Gynäkologin und im Rahmen einer Vorsorgeuntersuchung klagt die Ihnen bekannte 48-jährige Patientin, Frau S., über folgende Beschwerden, die seit etwa 8 Monaten bestünden: Hitzewallungen, Schweißausbrüche, Schlafstörungen, starke Stimmungs- und Antriebsschwankungen sowie Niedergeschlagenheit. Die Beschwerden seien zumindest so stark, dass sich Frau S. auch gelegentlich in ihrer Berufsausübung beeinträchtigt fühlt. Eine Freundin, die unter ähnlichen Beschwerden gelitten hätte, habe ihr berichtet, dass alle Beschwerden unter Hormontabletten wie „weggepustet“ waren. Frau S. fragt, ob eine derartige Behandlung auch für sie in Frage käme. Sie hätte gegen eine Tabletteneinnahme nichts einzuwenden. Frau S. ist normalgewichtig (BMI 23), sportlich (spielt Tennis), raucht gelegentlich, ist freiberuflich tätig (Inhaberin einer kleinen Boutique) und verheiratet. Zwei gesunde Töchter (23 und 25 Jahre) sind bereits aus dem Haus. Bei der Mutter der Patientin wurde vor zwei Jahren ein Mammakarzinom diagnostiziert. Frau S. hat zwei ältere, gesunde Schwestern. Die allgemeine und gynäkologische Anamnese der Patientin ist ansonsten unauffällig, ebenso die Befunde der körperlichen Untersuchung. In der letzten Zeit sei die Periode unregelmäßig.

Schätzen Sie bitte das Risiko einer Hormonerersatz-Therapie für die Entwicklung von Brustkrebs unter Zuhilfenahme einer Fall-Kontroll- und einer Kohortenstudie ein und berücksichtigen Sie dabei auch das genetische Risiko.

Folgt man der Abbildung 10a.1, lassen sich Studien grundsätzlich in experimentelle und beobachtende Designs unterteilen. Unter den beobachtenden Studien kommen im Hinblick auf die Abklärung ätiologischer Fragestellungen insbesondere der Fall-Kontroll- und der Kohortenstudie besondere Bedeutung zu. Beide Studienformen sollen im Folgenden unter Berücksichtigung des dargestellten klinischen Szenarios und unter klinisch-epidemiologischen Gesichtspunkten beschrieben und bewertet und im Anschluss auch unter genetischen Gesichtspunkten abgehandelt werden.

10a.2 Fall-Kontroll-Studie

Das Design der Fall-Kontroll-Studie ist in Abbildung 10a.2 grafisch dargestellt. Bei diesem Studientyp wird einer Gruppe bereits erkrankter Menschen (Fälle) eine Gruppe von Personen zugeordnet, die nicht unter der interessierenden Erkrankung leiden (Kontrollen). Der Untersucher blickt dann zeitlich zurück (retrospektiv), um die Häufigkeit von Risikofaktoren in beiden Gruppen festzustellen. Da in Fall-Kontroll-Studien aufgrund des retrospektiven Designs keine Inzidenzen und damit keine echten Risiken berechnet werden können, kann als Maß der Assoziation auch kein relatives Risiko angegeben werden. Letzteres wird in Fall-Kontroll-Studien

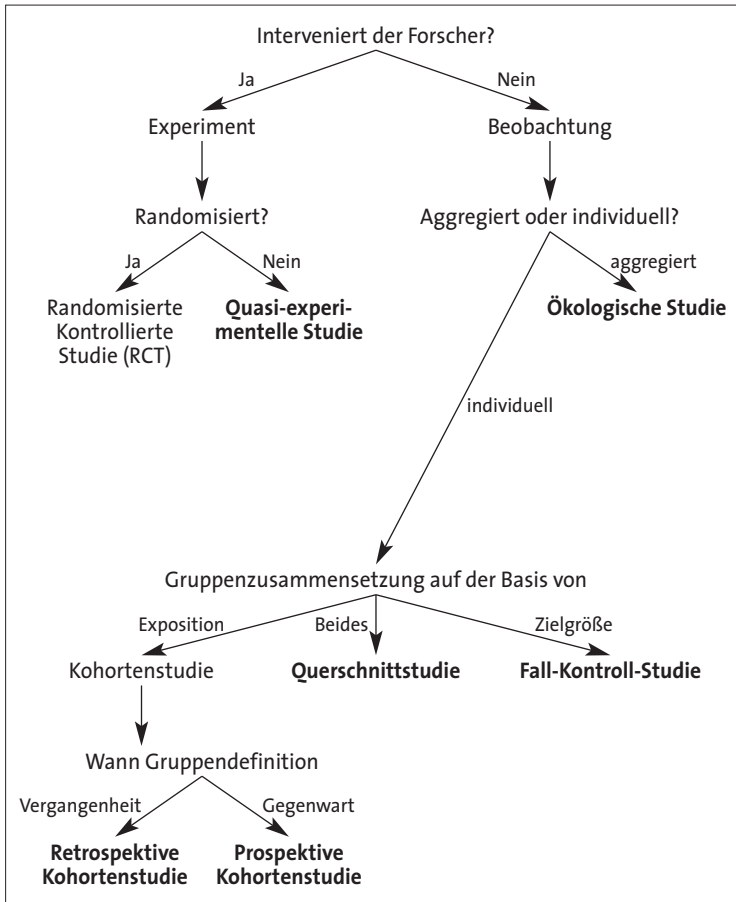


Abb. 10a.1: Algorithmus der Studientypen in klinischer Forschung [1]

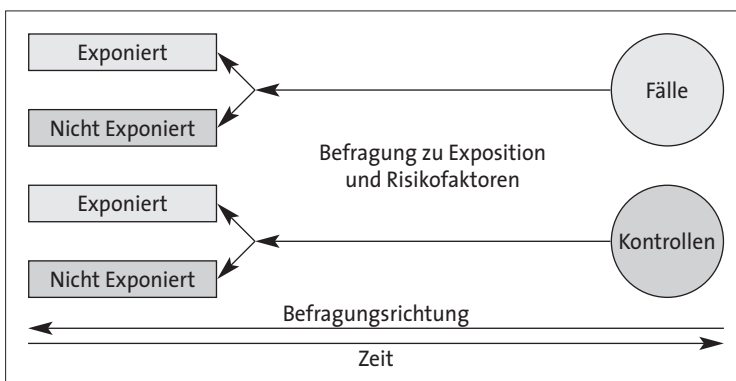


Abb. 10a.2: Ablaufschema der Fall-Kontroll-Studie

über das Chancenverhältnis (Odds-Ratio) approximiert, das sich in einer Vierfeldertafel aus dem Quotienten axd/bxc berechnet. Für gewöhnlich wird zusätzlich ein 95%-Konfidenzintervall als Maß der Stabilität angegeben.

10a.2.1 Vorteile der Fall-Kontroll-Studie

Die Fall-Kontroll-Studie stellt ein effizientes und ökonomisches Verfahren zur Untersuchung von Risikofaktoren dar. Da in der Fallgruppe die Erkrankung bereits eingetreten ist, muss nicht – wie in der prospektiven Kohortenstudie – das Auftreten von Neuerkrankungen abgewartet werden. Die Untersucher können Fälle auch unabhängig von der natürlichen Häufigkeit der Erkrankung rekrutieren, sodass sich dieses Verfahren insbesondere für seltene Erkrankungen eignet. Die Fall-Kontroll-Studie erlaubt außerdem die Untersuchung mehrerer Einflussfaktoren für eine Erkrankung.

10a.2.2 Nachteile der Fall-Kontroll-Studie

Ein kritischer Punkt in Fall-Kontroll-Studien ist die Auswahl einer geeigneten Kontrollgruppe. Diese sollte im Idealfall die Expositionsverteilung der Studiengrundgesamtheit, aus der die Fälle stammen, widerspiegeln. Kontrollen müssen unabhängig von ihrer Exposition ausgewählt werden und können z.B. eine Zufallsstichprobe aus der Studienpopulation darstellen oder im Krankenhaus oder aus der Nachbarschaft der Fälle rekrutiert werden. Kommt es zu Verzerrungen in der Auswahl der Kontrollen, weil diese beispielsweise nicht repräsentativ für die Expositionsverteilung sind oder in ihrem Risikoprofil den Fällen zu sehr ähneln, kann ein systematischer Fehler entstehen (Selection-Bias), der das Ergebnis differenziell verzerrt.

Fall-Kontroll-Studien sind außerdem insbesondere anfällig für systematische Fehler im Sinne des Informations-Bias (Recall-Bias), da erkrankte Personen sich meist differenziell besser an Exposition erinnern und es so zu einer Überschätzung des Effekts kommen kann. Im Gegensatz zu Kohortenstudien kann in Fall-Kontroll-Studien in der Regel nicht die zeitliche Abfolge des Auftretens von Exposition und Erkrankung dargestellt werden, sodass die Interpretation im Hinblick auf Kausalzusammenhänge eingeschränkt ist.

Um den Einfluss von Confoundern (Störgrößen) zu begrenzen, stehen verschiedene Möglichkeiten auf der Ebene des Studiendesigns und der Datenanalyse zur Verfügung. Im Hinblick auf das Design kann die zu untersuchende Population z.B. im Hinblick auf das Geschlecht auf Männer oder Frauen eingeschränkt werden (Restriktion). Geschlecht kann dann nicht mehr als Störgröße wirken, allerdings kann dieser Faktor auch nicht mehr als Risikofaktor analysiert werden und Aussagen für die nicht untersuchte Gruppe sind in der Regel auch nicht möglich. In Fall-Kontroll-Studien werden Fälle und Kontrolle häufig im Hinblick auf wichtige Störgrößen zugeordnet (matching). Durch diese Zuordnung, die auf individueller oder Gruppenbasis (Frequency matching) geschehen kann, werden Fälle und Kontrollen für diese Faktoren gleichgemacht, sodass diese wiederum nicht als Störgrößen fungieren können. Grundsätzlich müssen solche Studien dann auch gematcht analysiert werden, wobei wiederum Matching-Faktoren nicht als Risikofaktoren analysiert werden können.

Eine Gefahr besteht im so genannten Overmatching, d.h. es werden zu viele Faktoren als Matching-Faktoren ausgewählt, bzw. solche, die mit den zu untersuchenden Risikofaktoren verbunden sind, sodass Fälle und Kontrollen auch im Hinblick auf die interessierenden Risikofaktoren gleichgemacht werden, sodass eine Analyse diesbezüglich nicht

mehr möglich ist und ein tatsächlicher Effekt übersehen werden kann. Neben der Stratifikation bieten z.B. multiple Regressionsverfahren die Möglichkeit, den Einfluss von Störgrößen in der Analyse zu kontrollieren.

10a.3 Kohortenstudien

In prospektiven Kohortenstudien werden Gruppen von Personen über die Zeit hinweg beobachtet, die unterschiedlichen Expositionen ausgesetzt sind, wobei per definitionem zu Beginn der Studie die interessierende Erkrankung noch nicht aufgetreten ist. Es wird die Inzidenz im Sinne eines absoluten Risikos in beiden Gruppen bestimmt (s. Abb. 10a.3). Als Assoziationsmaß lassen sich dann relative Risiken berechnen und auch hier 95%-Konfidenzintervalle als Maß der Stabilität angeben.

10a.3.1 Vorteile von Kohortenstudien

Kohortenstudien eignen sich insbesondere bei seltenen Expositionen, da zu Studienbeginn eine Gruppe Exponierter gesammelt und einer Gruppe Nicht-Exponierter gegenübergestellt wird. Kohortenstudien erlauben die Untersuchung verschiedener Effekte (Outcomes) durch eine Exposition. Durch die in die Zukunft gerichtete Untersuchung und die Erkrankungsfreiheit zu Beginn ist

auch eine klare zeitliche Zuordnung von Exposition und Erkrankung im Sinne eines kausalen Aufeinanderfolgens möglich. Kohortenstudien erlauben eine valide und von Einflüssen der Probanden weitestgehend unabhängige Expositionserfassung.

10a.3.2 Nachteile von Kohortenstudien

Auf der anderen Seite sind Kohortenstudien deutlich zeit- und kostenintensiver als Fall-Kontroll-Studien. Sie sind für seltene Erkrankungen oder solche mit langer Latenzzeit weniger geeignet. Dieser Studientyp kann durch Studienabbrecher gefährdet werden, die sich qualitativ und quantitativ zwischen den beobachteten Gruppen unterscheiden können und so zu systematischen Verzerrungen beitragen.

Die Charakteristika von Kohorten- und Fall-Kontroll-Studien sind noch einmal in der folgenden Tabelle zusammengefasst.

10a.4 Die Evidenz

Zur Beantwortung der eingangs geschilderten klinischen Frage sollen zwei Studien herangezogen werden: die von Stanford et al. 1995 publizierte Fall-Kontroll-Studie [3] und die von Tamimi und Mitarbeitern publizierten Ergebnisse der Nurses' Health Study [4], die eine prospektive Kohortenstudie darstellt.

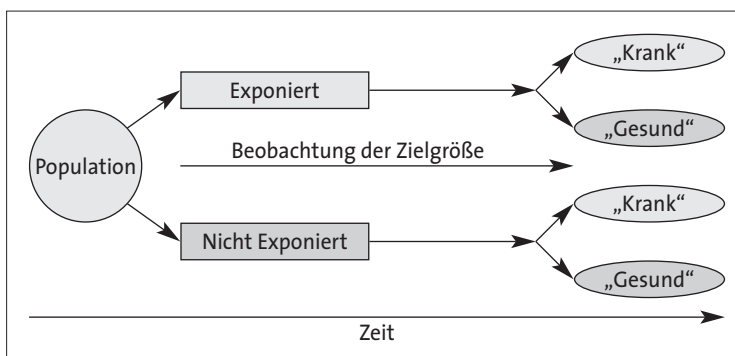


Abb. 10a.3: Ablaufschema der Kohortenstudie

Tab. 10a.1: Vergleichende Darstellung von Charakteristika der Fall-Kontroll- und Kohortenstudie [2]

Fall-Kontroll-Studie	Kohortenstudie
Population unter Risiko oft nicht definiert	beginnt mit einer definierten Population unter Risiko
Fälle durch den Untersucher aus einem zur Verfügung stehenden Patientenbestand selektiert	Fälle nicht selektiert, aber durch kontinuierliche Beobachtung ermittelt (Erfassung aller Fälle angestrebt)
Kontrollen durch den Untersucher selektiert (oft ähnlich den Fällen)	Kontrollen, Vergleichsgruppe (d.h. Nicht-Fälle), nichtselektiert – entwickeln sich natürlich
Exposition wird erfasst, rekonstruiert oder aus der Erinnerung wiedergegeben, zeitlich nach der Entwicklung der Krankheit	Exposition wird vor der Entwicklung der Krankheit erfasst
Risiko oder Inzidenz der Krankheit sind nicht aus der Studie bestimmbar: das relative Expositionsrisiko kann durch das Odds-Ratio geschätzt werden	Risiko oder Inzidenz der Krankheit und relatives Risiko sind direkt angebar

10a.4.1 Bewertung der Evidenz

Für beide Studientypen liegen checklistenartige Fragenkataloge vor, die eine strukturierte kritische methodische Bewertung und Ergebnisdarstellung erlauben. In den Tabellen 10a.2 und 10a.3 haben wir eine Übersetzung und Modifikation der Checklisten des Scottish Intercollegiate Guideline Networks (SIGN) [5] dargestellt.

10a.5 Die Fall-Kontroll-Studie

Die populationsbezogene Fall-Kontroll-Studie von Stanford et al. [3] untersucht den Einfluss der kombinierten menopausalen Hormonersatz-Therapie (Östrogen und Progesteron) auf das Brustkrebsrisiko. Als Fälle wurden 537 Frauen im Alter von 50 bis 64 Jahren eingeschlossen, die zwischen dem 1. Januar 1988 und dem 30. Juli 1990 mit einer Brustkrebs-Primärdiagnose in einem populationsbezogenen Krebsregister erfasst wurden, das 13 Bezirke des Staates Washington, USA, erfasst. Demgegenüber wurden 492 zufällig ausgewählte Kontrollprobandinnen gestellt, die aus derselben Region mittels zufällig generierter Telefonnummern identifiziert wur-

den und in 5-Jahres-Altersgruppen der Fallgruppe zugeordnet (gematcht) wurden.

Im Ergebnis hatten 57,6% der Fälle und 61,0% der Kontrollen menopausale Hormone eingenommen. Die Frauen, die jemals Kombinationspräparate eingenommen hatten, repräsentieren 21,5% der Fälle und 21,3% der Kontrollen und hatten danach kein erhöhtes Brustkrebsrisiko (OR 0,9; 0,7–1,3). Frauen, die Kombinationspräparate über mindestens acht Jahre eingenommen hatten, hatten verglichen mit solchen, die keine Hormonpräparate eingenommen hatten, ein geringeres Brustkrebsrisiko (OR 0,4; 0,2–1,0). Die Ergebnisse basieren auf Regressionsanalysen, die wesentliche Confounder wie Alter, Zeitpunkt der Menarche, Schwangerschaften, Familienanamnese bezüglich Brustkrebs sowie andere medizinische und Lebensstilfaktoren berücksichtigten.

10a.5.1 Sind die Ergebnisse der Studie valide?

Die Fragestellung ist sicherlich angemessen und bezogen auf die Exposition, die Zielgröße und die untersuchte Population klar fokussiert. Die Fälle stammen aus einem popu-

Tab. 10a.2: Kritische Bewertung von Fall-Kontroll-Studien (modifiziert nach SIGN [5])

Teil 1: Interne Validität	
1.1	Ist die Fragestellung angemessen und klar fokussiert?
Auswahl der Studienpopulation	
1.2	Stammen die Fälle und Kontrollen aus vergleichbaren Grundgesamtheiten?
1.3	Wurden für Fälle und Kontrollen die gleichen Ausschlusskriterien angewandt?
1.4	Wie hoch war die Teilnahmequote jeder einzelnen Teilnehmergruppe (Fälle und Kontrollen)?
1.5	Gibt es einen Vergleich von Teilnehmern und Nicht-Teilnehmern, um ihre Ähnlichkeiten und Unterschiede zu ermitteln?
1.6	Sind die Fälle klar definiert und von den Kontrollen abgegrenzt?
1.7	Ist sichergestellt, dass die Kontrollen keine Fälle sind?
Erfassung	
1.8	Wurden Maßnahmen ergriffen, um zu verhindern, dass die Kenntnis über die primäre Exposition die Definition des Fallstatus beeinflusst?
1.9	Wurde die Exposition auf eine standardisierte, valide und reliable Weise gemessen?
Confounder/Störgrößen	
1.10	Wurden die wichtigsten potenziellen Confounder identifiziert und angemessen im Studiendesign und in der Analyse berücksichtigt?
Statistische Analyse	
1.11	Sind Konfidenzintervalle berechnet worden?
Teil 2: Gesamtbeurteilung der Studie	
2.1	Wie gut wurde die Studie durchgeführt, um das Risiko von Bias oder Confounder zu minimieren, und um die kausale Beziehung zwischen Exposition und Zielgröße zu begründen?
2.2	Unter Berücksichtigung klinischer Überlegungen, Ihrer Evaluation der Methodik und der statistischen Power der Studie sind Sie sich sicher, dass der Gesamteffekt auf der beobachteten Exposition beruht?
Wenn der Studienbericht eine Evaluation oder ein Vergleich von diagnostischen Tests ist, vervollständigen Sie bitte die Checkliste der diagnostischen Studien, bevor Sie den nächsten Teil beenden.	
Teil 3: Beschreibung der Studie	
3.1	Welche Exposition oder welcher prognostische Faktor wird in dieser Studie evaluiert?
3.2	Welche Zielgrößen werden beschrieben?
3.3	Wie viele Patienten nehmen an der Studie teil? (Gesamtzahl und Anzahl in jeder Studiengruppe)
3.4	Welches sind die Merkmale der Studienpopulation? (z.B. Alter, Geschlecht, Krankheitsbild der Population, Krankheitsprävalenz)
3.5	Welches sind die Merkmale der Studienumgebung? (z.B. Land, Stadt, Krankenhauspatienten oder ambulant behandelt, Allgemeinarztpraxen, Umgebung)
3.6	Gibt es spezielle Fragen, die durch die Studie aufgeworfen wurden? Geben Sie einige generelle Kommentare zu den Studienergebnissen und ihrer Bedeutung.
3.7	Hauptergebnisse (Maße der Assoziation und Stabilität)
3.8	Ableitbare Empfehlungen

Tab. 10a.3: Kritische Bewertung von Kohortenstudien (modifiziert nach SIGN [5])

Teil 1: Interne Validität	
1.1	Ist die Fragestellung geeignet und klar fokussiert?
Auswahl der Studienpopulation	
1.2	Stammen die beiden untersuchten Gruppen von Grundgesamtheiten ab, die bis auf den untersuchten Faktor vergleichbar sind?
1.3	Wie sind die Teilnahmequoten in beiden Gruppen?
1.4	Wurde die Wahrscheinlichkeit abgeschätzt, dass geeignete Probanden bereits zu Beginn der Studie die im Rahmen der Studie zu untersuchende Zielgröße aufwiesen? Wurde dies in der Analyse berücksichtigt?
1.5	Wie hoch war der Anteil von Personen oder Gruppen in jedem Studienarm, die die Studie abbrachen?
1.6	Wurde überprüft, ob sich Personen, die bis zum Ende in der Studie blieben, von denen, die vorzeitig ausgeschieden sind, hinsichtlich des Expositionsstatus' unterschieden?
Erfassung	
1.7	Sind die Zielgrößen klar definiert?
1.8	Wurde die Erfassung der Zielgröße ohne Kenntnis des Expositionsstatus durchgeführt?
1.9	Wenn eine Verblindung nicht möglich war: Gibt es Hinweise, dass die Kenntnis über den Expositionsstatus die Zielgrößenerfassung beeinflusst hat?
1.10	Sind die Methoden zur Expositionserfassung reliabel?
1.11	Werden externe Nachweise für die Validität und Reproduzierbarkeit der Zielgrößenerfassung erbracht?
1.12	Sind der Expositionsstatus bzw. die prognostischen Faktoren mehr als einmal erfasst worden?
Confounder/Störgrößen	
1.13	Wurden die wichtigsten potenziellen Confounder identifiziert und angemessen im Studiendesign und in der Analyse berücksichtigt?
Statistische Analyse	
1.14	Sind Konfidenzintervalle berechnet wurden?
1.15	Ist ein Maß für den Goodness of fit für die verwendeten multivariaten Modelle angegeben?
1.16	Ist eine Korrektur für das multiple statistische Testen durchgeführt wurden?
Teil 2: Gesamtbeurteilung der Studie	
2.1	Wie geeignet war die Studie, das Risiko von Verzerrungen oder den Einfluss von Confoundern zu minimieren und eine kausale Beziehung zwischen Intervention/Exposition und Effekt nachzuweisen?
2.2	Unter Berücksichtigung klinischer Überlegungen, Ihrer Evaluation der Methodik und der statistischen Power der Studie sind Sie sich sicher, dass der Gesamteffekt auf der Studienintervention beruht?
Wenn der Studienbericht eine Evaluation oder ein Vergleich von diagnostischen Tests ist, vervollständigen Sie bitte die Checkliste der diagnostischen Studien, bevor Sie den nächsten Teil beenden	

Tab. 10a.3: Fortsetzung

Teil 3: Beschreibung der Studie	
3.1	Welche Exposition oder welcher prognostische Faktor wurde in der Studie evaluiert? Wurden An- und Abwesenheit dieser Faktoren oder verschiedene Ausprägungsgrade miteinander verglichen?
3.2	Welche Zielgrößen wurden erfasst und wie lange war die Beobachtungszeit?
3.3	Wie viele Patienten nahmen an der Studie teil? (Gesamtzahl und Anzahl in jeder Studiengruppe)
3.4	Welches sind die Merkmale der Studienpopulation? (z.B. Alter, Geschlecht, Krankheitsbild der Population, Krankheitsprävalenz)
3.5	Welches sind die Merkmale der Studienumgebung? (z.B. Land, Stadt, Krankenhauspatienten oder ambulant behandelt, Allgemeinarztpraxen, Umgebung)
3.6	Gibt es spezielle Fragen, die durch die Studie aufgeworfen wurden? Geben Sie einige generelle Kommentare zu den Studienergebnissen und ihrer Bedeutung.
3.7	Hauptergebnisse (Maße der Assoziation und Stabilität)
3.8	Ableitbare Empfehlungen

lationsbezogenen Krebsregister und wurden auf Einwohnerinnen des Landkreises King County des Staates Washington beschränkt. Kontrollen der gleichen Altersgruppe wurden mittels zufällig ausgewählter Telefonnummern identifiziert und auf Einwohnerinnen des gleichen Landkreises beschränkt, sodass es keine Hinweise darauf gibt, dass Fälle und Kontrollen nicht die gleiche Grundgesamtheit abbilden. Die Auswahl von Fällen und Kontrollen folgte in Bezug auf Altersgruppen, weitere Ausschlusskriterien sind nicht genannt.

Die Nichtteilnehmerquote lag bei 18,6% in der Fall- und bei 27,0% in der Kontrollgruppe. Unter den Fällen lagen die Gründe für die Nichtteilnahme überwiegend in einer allgemeinen Verweigerung (12,6%), weniger in logistischen (1,5%) oder Kommunikationsproblemen (0,6%). 3,9% der potenziellen Fälle waren bereits verstorben. Gründe für die Nichtteilnahme der Kontrollprobandinnen sind nicht explizit benannt. Die unterschiedlichen Teilnahmequoten bergen das Risiko des Selektionsbias, d.h. der im Hinblick auf Exposition und Fallstatus differenziellen Teilnahme, die zu einer systematischen Verzerrung des Effektschätzers führen

kann. Informationen über die Hormoneinnahme bei Nichtteilnehmern lagen nicht vor.

In der Arbeit werden wesentliche Charakteristika und Confounder vergleichend zwischen Fällen und Kontrollen in einer Tabelle aufgeführt. Dabei wird deutlich, dass Fälle gegenüber den Kontrollen besser ausgebildet waren, sowie häufiger kinderlos waren und eine spätere Erstschwangerschaft, einen höheren Body-Mass-Index und in einem höheren Anteil eine positive Brustkrebsfamilienanamnese hatten.

Die Falldefinition orientierte sich an histologisch abgesicherten inzidenten Fällen von invasivem oder in situ Brustkrebs, die in dem Krebsregister registriert wurden. Daten zur Erfassungsqualität (Vollzähligkeit und Vollständigkeit) des Krebsregisters werden nicht angegeben. Potenzielle Kontrollprobandinnen wurden bezüglich ihrer Brustkrebsanamnese befragt und insgesamt 53 Frauen mit positiver Anamnese wurden aus der Kontrollgruppe ausgeschlossen.

Es ist davon auszugehen, dass im Ablauf der Studie zunächst der Fallstatus gesichert und anschließend die Exposition erfasst wurde, sodass die potenzielle Kenntnis über die

primäre Exposition kaum die Definition des Fallstatus beeinflusst haben dürfte. Die Exposition, d.h. die Hormonersatz-Therapie wurde in standardisierter Weise mittels eines strukturierten Fragebogens erhoben. Daten zur Validität oder Reliabilität der Expositionserfassung werden nicht gezeigt. Der Fragebogen wurde als Grundlage für ein persönliches Interview benutzt. Da davon auszugehen ist, dass die Interviewer den Fallstatus kannten, ist eine differenzielle Beeinflussung der Expositionserfassung durch den Interviewer nicht auszuschließen (Interviewer-Bias). Die Autoren haben die wesentlichen Störgrößen im Interview erfasst und angemessen in der Analyse im Sinne logistischer Regressionsanalysen berücksichtigt. Seitens des Studiendesigns wurden durch Restriktion und Matching die Faktoren Geschlecht und Alter berücksichtigt.

Nie auszuschließen ist, dass unbekannte und nicht erfasste Störgrößen das Ergebnis beeinflusst haben (residual confounding). Der qualitative und quantitative Einfluss diesbezüglich ist allerdings schwer abzuschätzen.

Im Hauptergebnis werden adjustierte Odds-Ratios und 95%-Konfidenzintervalle als Maße der Assoziation und Stabilität angegeben. Neben den genannten Einschränkungen gilt es zu berücksichtigen, dass Kombinationspräparate erst relativ kurz vor Beginn der Studie in größerem Umfang verschrieben wurden. Dies bedeutet, dass Latenzzeiten, wie sie für die Induktion von Krebserkrankungen bekannt sind, in dieser Studie nicht berücksichtigt werden konnten.

Eine wesentliche Schwäche in dem beobachtenden Design der Fall-Kontroll-Studie liegt darin, dass die Exposition nicht zufällig zugeteilt wurde. Im Praxisalltag wird die Indikationsstellung für eine Hormonersatz-Therapie das individuelle Risikoprofil berücksichtigen. Dies bedeutet, dass Frauen mit einem erhöhten Brustkrebsrisiko seltener eine Hormonersatz-Therapie erfahren

und dass im Praxisalltag die behandelte Gruppe Frauen umfassen wird, die ein selektiv günstigeres Risikoprofil bezüglich Brustkrebs aufweisen. Dies wird zu einer differenziellen Unterschätzung des Effekts führen.

Fasst man die dargestellten Einschränkungen dieser Fall-Kontroll-Studie zusammen, so erscheint die Aussagekraft zur Assoziation zwischen Hormonersatz-Therapie und Brustkrebsrisiko sehr eingeschränkt. Es ist wahrscheinlich, dass vorhandene Effekte unterschätzt wurden.

10a.6 Die Kohortenstudie

Weniger störanfällig für systematische Fehler ist die prospektive Kohortenstudie. Der randomisierte, kontrollierte Versuch (RCT) kann als Sonderform der prospektiven Kohortenstudie aufgefasst werden. Klassischerweise wird die Fallzahl in einem RCT anhand der zu erwartenden Wirkung und nicht anhand der zu erwartenden Nebenwirkung kalkuliert. Deswegen erkennen RCTs normalerweise Nebenwirkungen nicht in statistisch auffälliger Weise. Formal gilt der RCT auch als Goldstandard für die Erkennung von Nebenwirkungen. Um vergleichsweise seltene Nebenwirkungen, wie eine Erhöhung der Brustkrebsinzidenz-Rate statistisch abgesichert erkennen zu können, bedarf es großer Kohortenstudien mit einer vergleichsweise langen Beobachtungsdauer. Eine derartige Studie stellt die Nurses' Health Study [4] dar. Für diese Analyse wurde ein Beobachtungszeitraum von 24 Jahren (1978–2002) ausgewertet. Initial wurden 35 474 post-menopausale Frauen eingeschlossen. Frauen, die im Verlauf der Beobachtungsperiode die Menopause erreichten, wurden sukzessive eingeschlossen, sodass im Jahre 2000 70 444 Frauen zu der Analyse beitrugen. Die Erhebung von Zielgrößen, Exposition und Confoundern erfolgte alle zwei Jahre über Fragebogen. Für diese Analyse wurde die post-me-

nopausale Hormonersatz-Therapie in den Kategorien „ausschließliche Östrogeneinnahme“, „ausschließliche Testosteroneinnahme“ und „kombinierte Östrogen-/Testosterontherapie“ unterteilt. Im Beobachtungszeitraum traten bei einer Anzahl von 1359323 Personenjahren insgesamt 4610 inzidente Brustkrebsfälle auf. Nach Ausschluss hysterektomierter Frauen ergab sich ein rund 2½-fach erhöhtes Brustkrebsrisiko (RR 2,48; 95% KI, 1,53–4,04) für Frauen, die Kombinationspräparate eingenommen hatten, gegenüber denjenigen, die keine Hormonersatz-Therapie erhalten hatten. Das Brustkrebsrisiko unter Kombinationstherapie war signifikant höher als das unter ausschließlicher Östrogentherapie. In der multivariaten Analyse der Gesamtgruppe finden sich gegenüber Frauen, die keine Hormonersatz-Therapie erhalten hatten, erhöhte Risiken für die Einnahme von Östrogen allein (RR 1,15, 95% KI, 1,05–1,27), die Kombinationsbehandlung aus Östrogen und Testosteron (RR 1,77, 95% KI, 1,22–2,56) und die alleinige Einnahme von Testosteron (RR 2,52, 95% KI, 0,80–7,94). Die stratifizierte Analyse für eine Einnahmedauer von unter bzw. über fünf Jahren zeigt höhere Risiken bei längerer Hormonersatz-Therapie.

10a.6.1 Sind die Ergebnisse der Studie valide?

Sicherlich ist die Studie geeignet, um die Fragestellung des Einflusses der Hormonersatz-Therapie auf das Brustkrebsrisiko zu untersuchen und ist in der Analyse auch klar bezüglich dreier Verabreichungsschemata fokussiert. Da in der Nurses' Health Study innerhalb der Gesamtkohorte nach Exposition stratifiziert wurde, ist davon auszugehen, dass die untersuchten Gruppen von einer Grundgesamtheit abstammen, die bis auf die untersuchte Exposition vergleichbar ist. Angaben zur Teilnahmequote sind dieser Publi-

kation nicht zu entnehmen. Frauen, die zum Zeitpunkt der Ausgangserhebung eine Tumordiagnose hatten, wurden von der weiteren Analyse ausgeschlossen. Ein Screening zur Früherkennung von Brustkrebs erfolgte zum Einschlusszeitpunkt nicht. Bezüglich der Quote an Studienabbrechern wird berichtet, dass sich im Jahr 2002 noch über 90% der Frauen in der Kohorte befanden. Genauere Angaben im Hinblick auf Abbruchgründe oder quantitativ wie qualitativ unterschiedliches Abbruchverhalten nach Exposition erfolgten nicht. Dementsprechend ist auch keine Aussage darüber möglich, ob sich die Studienabbrecherinnen hinsichtlich des Expositionsstatus' von den Teilnehmerinnen unterscheiden. Die Zielgrößenerfassung basiert auf der Selbstangabe einer Brustkrebsdiagnose. Diese Diagnosen wurden anhand der Durchsicht der Krankenakten bestätigt. Eine entsprechende positive Bestätigung erfolgt in über 99% der berichteten Brustkrebsfälle. Hinsichtlich der Frage, ob die Erfassung der Zielgröße ohne Kenntnis des Expositionsstatus durchgeführt wurde, sind keine Angaben gemacht. Die Kenntnis über den Expositionsstatus mag die Zielgrößenerfassung insofern beeinflusst haben, als dass es wahrscheinlicher ist, dass Frauen unter Hormonersatz-Therapie eher ein Brustkrebs-Screening erhalten (Detection-Bias). Dies würde allerdings nicht die aufgezeigten unterschiedlichen Risiken für unterschiedliche Hormonersatz-Therapien erklären. Die Expositionserfassung erfolgte über die Fragebögen nach Eigenangaben. Über die Beobachtungsperiode hinweg waren die Fragen nicht immer identisch, z.B. wurde die Kombinationstherapie aus Östrogen und Testosteron nicht immer explizit abgefragt. Es bestand zwar die Möglichkeit zur Angabe als Klartext, aber eine Missklassifikation bezüglich dieser Kombinationstherapie im Sinne einer Unterschätzung der Exposition ist nicht auszuschließen. Brustkrebsfälle wurden anhand der Krankenakten bestätigt, aber

weitere Angaben zur Validität und Reproduzierbarkeit dieser Zielgrößenerfassung sind nicht gemacht. Expositionstatus, prognostische Faktoren bzw. Confounder und die Zielgrößen wurden in zweijährlichem Abstand erfasst. Die wichtigsten Confounder, wie z.B. das Alter zum Zeitpunkt der Menopause, die Familienanamnese bezüglich Brustkrebs, die Eigenanamnese bezüglich gutartiger Brusterkrankungen, Body-Mass-Index, Gewichtsveränderungen, Alter zum Zeitpunkt der Menarche, Geburten und Alter zum Zeitpunkt der ersten Geburt, Alkoholkonsum und Einnahme anderer Hormonpräparate wurden erfasst und in der multivariaten Analyse berücksichtigt. Als Maße der Assoziation und Stabilität wurden relative Risiken und 95%-Konfidenzintervalle berichtet. Maße für den Goodness of fit sind allerdings nicht angegeben. Eine Korrektur für multiples statistisches Testen erfolgte nicht, erscheint aber auch nicht notwendig. Neben den schon genannten Einschränkungen ist erwähnenswert, dass die Kombinationstherapie aus Östrogen und Testosteron erst in den letzten Jahren des Beobachtungszeitraumes in relevanter Häufigkeit und zunehmend durchgeführt wurde. Möglicherweise entsteht hier ein Confounding durch entsprechende vorangegangene Therapien. Zusammenfassend müssen folgende methodische Aspekte kritisch gewürdigt werden: Detektionsbias, Missklassifikation der Exposition, Verlust von Teilnehmern während der Durchführung der Studie (Loss to follow-up) und ver-

ändertes Therapieverhalten im Verlauf der relativ langen Beobachtungsperiode. Dennoch zeigt die Studie einen eindeutigen und im Vergleich zu der zuvor besprochenen Fall-Kontroll-Studie deutlich valideren Hinweis auf einen Zusammenhang zwischen einer Hormonersatz-Therapie und dem Brustkrebsrisiko.

Zieht man zur Beurteilung der **Kausalität** die Kriterien von Bradford und Hill [6] heran, so konnte die prospektive Kohortenstudie die zeitliche Beziehung von Exposition und Erkrankung eindeutig aufzeigen. Die Induktion von Brustkrebs durch Hormoneinnahme scheint nach den bisherigen Erkenntnissen plausibel. Die Studienergebnisse sind konsistent mit vergleichbaren jüngeren und großen Kohortenstudien, stehen aber überwiegend aus methodischen Gründen im Widerspruch insbesondere zu älteren Fall-Kontroll-Studien. Eine Metaanalyse von 51 Studien bestätigte allerdings das erhöhte Brustkrebsrisiko [7].

Berücksichtigt man, dass die Brustkrebsentstehung – wie auch in dem Kausalitätsmodell von Rothman (s. Abb. 10a.4) dargestellt – ein multifaktorielles Geschehen ist und die Hormoneinnahme eine beitragende Ursache darstellt, so ist eine Erhöhung des Krebsrisikos um das 2½-fache durchaus relevant. Weiterhin konnte die Analyse zeigen, dass das Risiko bei längerer Einnahme noch einmal deutlich erhöht ist, was im Sinne einer Dosis-Wirkungsbeziehung interpretiert werden kann. Daten zur Reversibilität kön-

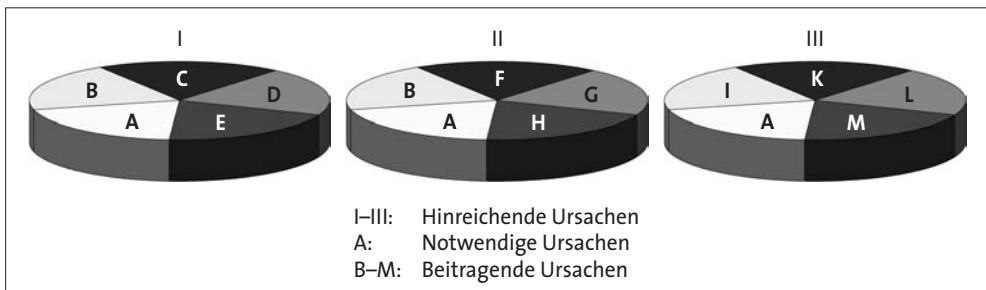


Abb. 10a.4: Kausalitätsmodell von K. Rothman [8]

Tab. 10a.4: Von Bradford erarbeiteter und durch Hill [9] weiterentwickelter Kriterienkatalog zur systematischen Prüfung auf Kausalität [10]

Zeitliche Beziehung:	Geht die Ursache der Wirkung voraus?
Plausibilität:	Stimmt die Assoziation mit anderen Erkenntnissen überein? (z.B. Wirkungsmechanismen, Ergebnissen aus Tierversuchen usw.)
Konsistenz:	Führten andere Studien zum gleichen Ergebnis?
Stärke:	Wie stark ist der Zusammenhang zwischen der Ursache und der Wirkung?
Dosis-Wirkungs-Beziehung:	Ist die Wirkung bei erhöhter Exposition gegenüber der vermuteten Ursache stärker?
Reversibilität:	Führt die Beseitigung der vermuteten Ursache zu einem geringeren Krankheitsrisiko?
Studienplan:	Basieren die Ergebnisse auf einem überzeugenden Studienplan?
Beurteilung der Evidenz für Kausalität:	Auf wie vielen Nachweisen/Beweisen beruhen die Ergebnisse?

nen aus der vorliegenden Studie nicht abgeleitet werden. Offensichtlich gelingt es mit dem Design der prospektiven Kohortenstudie nicht nur, die zeitliche Beziehung von Exposition und Zielgröße besser darzustellen, sondern durch die Ermittlung echter Risiken auch zu valideren Ergebnissen zu kommen.

10a.7 Fazit

Zusammenfassend haben die Ergebnisse der vorliegenden und anderer Kohortenstudien, insbesondere der Women's Health Initiative [4] sowie der One Million Women Study [11] auch in Deutschland zu einer Neubewertung der Hormonersatz-Therapie geführt [12]. Demnach soll die Indikationsstellung individuell und unter strenger Berücksichtigung des Brustkrebsrisiko-Profiles erfolgen. Die Hormontherapie sollte nur bei starken Beschwerden und so kurz wie möglich durchgeführt werden.

Die klinisch-epidemiologischen Ergebnisse haben auch Eingang in die Gerichtsbarkeit gefunden. In den USA haben rund 5 100 Frauen Klageschriften gegen die Firma Wyeth vorbereitet. In dem zweiten vor Gericht

zugelassenen Fall kam die Jury Anfang des Jahres 2006 zu dem Urteil, dass die Hormonersatz-Therapie zumindest teilweise für die Brustkrebserkrankung der Klägerin ursächlich war. Im weiteren Verlauf des Verfahrens wird geklärt werden, inwieweit die Firma dafür haftbar gemacht werden kann. Sollte dies der Fall sein, hat die Jury der Betroffenen und ihrem Ehemann eine Schadenssumme von insgesamt 1,5 Mio. \$ zugesprochen [13].

Das Bundesinstitut für Arzneimittel und Medizinprodukte hat 2003 angeordnet, die Produktinformation von Arzneimitteln zur Hormonersatz-Therapie um entsprechende Risikoangaben zu erweitern. Die Konsensempfehlungen zur Hormontherapie im Klimakterium und in der Postmenopause der Deutschen Gesellschaft für Gynäkologie und Geburtshilfe vom 5.5.2004 stellen fest, dass eine Hormontherapie nur bei bestehender Indikation eingesetzt und dass eine Nutzen-Risiko-Abwägung und Entscheidung zur Therapie gemeinsam mit der ratsuchenden Frau erfolgen soll. Nach Marion Caspers-Merk [14] scheint sich aktuell die Verschreibungspraxis allerdings wieder zu ändern: „Es zeichnet sich ab, dass wieder vermehrt Hormone auch dann eingesetzt werden, wenn es nicht zwingend notwendig ist. Über die Risi-

ken einer solchen Therapie kann niemand mehr im Unklaren sein. Ich fordere Ärztinnen und Ärzte auf, sich ihrer besonderen Verantwortung bewusst zu sein. Nach dem Stand der Wissenschaft kann es keinen Zweifel geben, dass die Hormonersatz-Therapie nur in besonderen Fällen und nur kurzzeitig eingesetzt werden sollte.“

Als niedergelassener Frauenarzt/Frauenärztin stehen Sie in einem klassischen Interessenskonflikt. Die Patientin wendet sich Hilfe suchend und konkret auch mit der Bitte um medikamentöse Therapie an Sie. Sie müssen fürchten, dass es das Arzt-Patienten-Verhältnis belastet, wenn Sie ihr diese Behandlung (aus guten Gründen) nicht zukommen lassen und dass Sie diese Patientin möglicherweise auch verlieren. Aus dem Blickwinkel des Praxisalltags können Ihnen zudem die Zusammenhänge zwischen Hormonersatz-Therapie und Brustkrebsrisiko auch nicht offenbar werden. Dennoch muss man unter Würdigung der vorhandenen Evidenz in dem hier skizzierten Fall entsprechend zurückhaltend mit einer Hormonersatz-Therapie sein.

Literatur

- [1] Katz DL (Ed) (2001) *Clinical epidemiology and evidence-based medicine. Fundamental principles of clinical reasoning and research.* Sage, Thousand Oaks
- [2] Fletcher RH et al. (1999) *Klinische Epidemiologie: Grundlagen und Anwendung.* Deutschsprachige Ausgabe. Ullstein Medical, Wiesbaden
- [3] Stanford JL et al., Combined estrogen and progestin hormone replacement therapy in relation to risk of breast cancer in middle-aged women. *JAMA.* (1995), 274, 137–142
- [4] Tamimi RM et al., Combined estrogen and testosterone use and risk of breast cancer in postmenopausal women. *Arch Intern Med.* (2006), 166, 1483–1489
- [5] Scottish Intercollegiate Guideline Network (SIGN). Critical appraisal: Notes and checklists. <http://www.sign.ac.uk/methodology/checklists.html> (24.11.06)
- [6] Hill AB, The environment and disease: association or causation? *Proc R Soc Med.* (1965), 58, 295–300
- [7] Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer. *Lancet.* (1997), 350, 1047–1059
- [8] Rothman KJ, *Causes.* *Am J Epidemiol.* (1976), 104, 587–592
- [9] Hill AB, The environment and disease: association or causation? *Proc R Soc Med.* (1965), 58, 295–300
- [10] Beaglehole R et al. (1993) *Basic epidemiology.* WHO, Geneva
- [11] Beral V, Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet.* (2003), 362, 419–427
- [12] Haas JS et al., Changes in the use of postmenopausal hormone therapy after the publication of clinical trial results. *Ann Intern Med.* (2004), 140, 184–188
- [13] The Associated Press. Woman wins first phase of cancer trial. http://www.quote.com/qc/news/story.aspx?symbols=APONLINE:100&story=200610042327_APO_83958363 (24.11.06)
- [14] Caspers-Merk M. Studie belegt Verschwendung und Risiken bei der Verordnung von Medikamenten. Pressemitteilung vom 20.6.2006. http://www.bmg.bund.de/nn_892718/DE/Presse/Pressemitteilungen/Presse-2-2006/pm-20-6-06,param=.html (24.11.06)

10b Genetische Studien

Inke R. König, Andreas Ziegler

Zum Verständnis der Ätiologie einer Erkrankung hat die Genetik in den vergangenen Jahrzehnten vermehrt beigetragen. So werden in der Datenbank „Online Mendelian Inheritance in Man“ [1] derzeit mehr als 1 800 Phänotyp-Beschreibungen gelistet, deren molekulare Grundlage bekannt ist. Diese Aufklärung genetischer Ursachen wurde unterstützt durch die Disziplin „Genetische Epidemiologie“, die sich spezifisch damit beschäftigt, genetische und Umweltursachen bei der Entstehung und Verbreitung von Erkrankungen zu untersuchen [2].

Hierbei werden verschiedene Studiendesigns verwendet, die sich zum einen in dem zugrunde liegenden Prinzip unterscheiden. Ein Teil der Designs hat nämlich zum Ziel, eine Kopplung zwischen einem genetischen Marker und einer Erkrankung zu entdecken [3]. Hierbei wird überprüft, ob der Marker gemeinsam mit der Erkrankung in einer Familie vererbt wird. Da hierfür die möglicherweise schwierige Rekrutierung von Familien erforderlich ist, haben Studien an Popularität gewonnen, die statt nach einer Kopplung nach genetischer Assoziation suchen. Von genetischer Assoziation wird dann gesprochen, wenn ein Allel an einem genetischen Marker bei erkrankten Personen häufiger vorkommt als bei gesunden. Damit ist die Rekrutierung von Familien nicht unbedingt erforderlich, und es können Studiendesigns eingesetzt werden, die auch in der klassischen Epidemiologie Anwendung finden. Besonders verbreitet sind Fall-Kontroll-Studien, bei denen in Analogie zur klassischen Epidemiologie das Allel am genetischen Marker als Risikofaktor betrachtet wird.

Eine zweite Unterscheidung der verwendeten Studiendesigns liegt in der Auswahl der untersuchten genetischen Marker. Bei einem systematischen Screening des gesamten menschlichen Genoms wird ohne vorherige Kenntnis möglicherweise relevanter Gene eine Vielzahl von Markern, über das Genom verteilt, analysiert. Im Vergleich dazu werden bei der Untersuchung von Kandidatengen spezifisch Marker in solchen Genen eingeschlossen, die aufgrund ihrer Funktion oder Position biologisch relevant sein könnten. In diesem Beitrag liegt der Fokus auf diesen Studien, die Prinzipien lassen sich jedoch problemlos übertragen. Im Folgenden beschreiben wir Kriterien, anhand derer eine solche genetische Assoziationsstudie bewertet werden kann. Zur Übersicht sind diese in Tabelle 10b.1 aufgeführt.

Szenario

Im Fall Ihrer Patientin stellt sich Ihnen aufgrund der Diagnose bei ihrer Mutter die Frage, ob eine genetische Prädisposition für ein Mammakarzinom besteht. Ihnen ist bekannt, dass einige Kandidatengene in der Literatur postuliert wurden, so zum Beispiel eine Variante im Methylenetetrahydrofolatreduktase (MTHFR)-Gen [4]. Sie möchten die Evidenz für dieses Kandidatengen beurteilen.

10b.1 Die Evidenz bewerten – Kritische Bewertung einer genetisch-epidemiologischen Assoziationsstudie

10b.1.1 Sind die Ergebnisse der Studie glaubwürdig?

Um die Glaubwürdigkeit einer genetisch-epidemiologischen Fall-Kontroll-Studie zu beurteilen, ist zunächst wichtig zu verstehen, welches die Ursachen für die Beobachtung einer genetischen Assoziation sein können. Im einfachsten Fall (vgl. Abb. 10b.1, oben) ist das Allel des genetischen Markers, der untersucht wird, ursächlich für die Erkrankung verantwortlich. In diesem Fall spricht man von einer direkten Assoziation. Bei einer indirekten Assoziation hingegen (vgl. Abb. 10b.1, Mitte) wird eine Assoziation zwischen dem Allel und der Erkrankung beobachtet, weil der untersuchte Marker in der Nähe eines anderen genetischen Locus liegt, der sei-

nerseits die Erkrankung verursacht. Diese Nähe zwischen zwei genetischen Loci wird als Kopplungsungleichgewicht bezeichnet.

Als dritte Möglichkeit kann eine beobachtete Assoziation, wie bei klassischen epidemiologischen Studien, durch Confounding verursacht sein. Entsprechend kommt auch bei diesen Studien der geeigneten Auswahl von Fällen und Kontrollen eine besondere Bedeutung zu.

Abhängig vom Ziel der Studie unterscheiden wir hier prinzipiell zwei Möglichkeiten. Das erste Ziel einer Assoziationsstudie ist häufig, überhaupt eine genetische Variante zu identifizieren, die mit der interessierenden Erkrankung assoziiert ist. Dafür können Probanden so ausgewählt werden, dass der genetische Effekt maximiert wird. Möglichkeiten hierfür sind in der Literatur beschrieben [2]. Im zweiten Schritt steht dann oftmals die genaue Schätzung des genetischen Effekts im Vordergrund. Hierfür ist, wenn auf die Untersuchung von Familien verzichtet

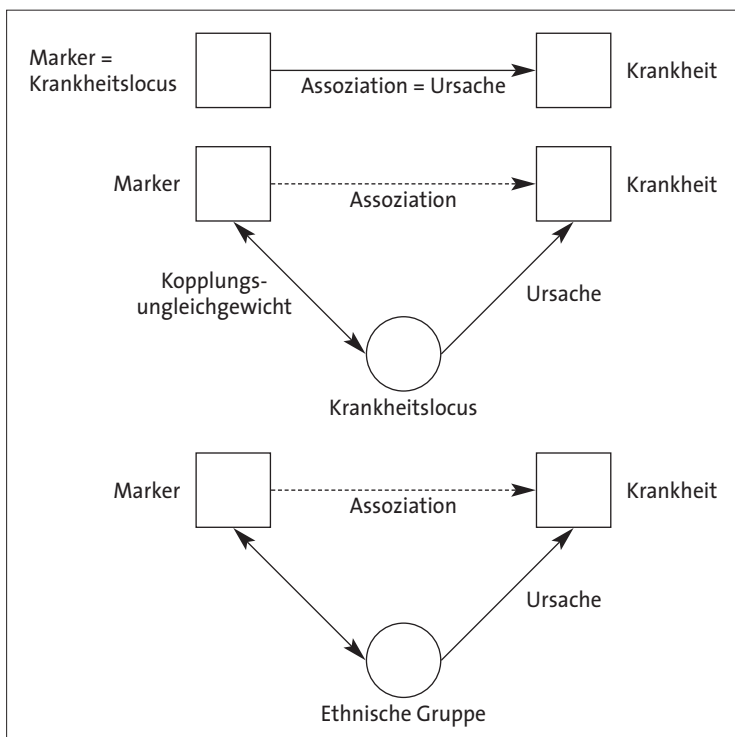


Abb. 10b.1: Ursachen für die Beobachtung einer genetischen Assoziation

Tab. 10b.1: Kriterien zur Beurteilung der Methodik einer genetischen Fall-Kontroll-Studie zur Ätiologie

1. Sind die Ergebnisse glaubwürdig?
<ul style="list-style-type: none"> • Wurden Fälle und Kontrollen nach einer begründeten Rationale ausgewählt? • Berücksichtigt die Studie wichtige Confounder durch Matching oder in der Auswertung? • Wurde Populationsstratifikation ausgeschlossen oder in der Auswertung berücksichtigt? • Ist die Genotypisierung verblindet durchgeführt worden? • Gab es in einer Gruppe Abweichungen vom Hardy-Weinberg-Gleichgewicht? Wenn ja, ist dies plausibel und wurde dies in der Auswertung berücksichtigt? • Ist der Phänotyp standardisiert erhoben? • Sind die Ergebnisse in einer unabhängigen Stichprobe oder in Funktionsstudien validiert?
2. Wie lauten die Studienergebnisse?
<ul style="list-style-type: none"> • Werden angemessene epidemiologische Maßzahlen mit Konfidenzintervallen berichtet? • Berücksichtigt die Interpretation die Zusammensetzung der Stichprobe? • Sind das Kandidatengen plausibel und die Marker sinnvoll ausgewählt? • Wie viele Marker in wie vielen Genen wurden zu wie vielen Phänotypen analysiert?

werden soll, die Rekrutierung von bevölkerungsbasierten Kontrollpersonen oder von Kohorten erforderlich. Bei beiden Vorgehensweisen müssen, wie bei klassischen epidemiologischen Studien auch, wichtige Confounder durch Matching zwischen Fällen und Kontrollen oder durch multivariate Auswertung berücksichtigt werden. Das bedeutet, dass für die Beurteilung der Glaubwürdigkeit einer Studie zum einen die Rationale für die Auswahl von Fällen und Kontrollen beschrieben und zum anderen eine adäquate Berücksichtigung der Confounder erfolgt sein muss.

In der vorliegenden Studie [4] wurden die Fälle, Frauen im Alter zwischen 45 und 75 Jahren mit inzidentem Brustkrebs, über Register identifiziert. Als Kontrollen wurden Frauen im gleichen Alter aus einer laufenden Kohortenstudie gezogen. Mögliche Confounder wurden in der multivariaten Auswertung berücksichtigt.

Neben den typischerweise relevanten Confoundern ist bei genetischen Studien spezifisch das Confounding durch Ethnizität, auch als Populations-Stratifikation bezeichnet, zu beachten (Abb. 10b.1, unten). Dies kann dadurch entstehen, dass die untersuchte Stichprobe hinsichtlich der ethnischen Zugehörigkeit heterogen ist. Wenn

sich dann sowohl die Prävalenz der Erkrankung als auch die Allelfrequenz des untersuchten genetischen Markers in den ethnischen Gruppen unterscheidet, kann möglicherweise eine Assoziation beobachtet werden, die nicht kausal auf den genetischen Marker zurückzuführen ist. Um dies zu vermeiden, werden die Fall- und Kontrollgruppen häufig nach Ethnizität gematcht. Da hierdurch aber subtile Unterschiede nicht erfasst werden können, ist ein Test auf das Vorliegen von Populations-Stratifikation zu empfehlen und, falls erforderlich, die statistische Korrektur dafür.

In der Studie waren Frauen aus mehr als fünf ethnischen Gruppen vertreten. In der Auswertung wurde für die Zugehörigkeit zu diesen Gruppen adjustiert, aber nicht auf Populations-Stratifikation innerhalb dieser Gruppen getestet.

Ein weiterer Aspekt bei der Beurteilung der Glaubwürdigkeit ist die Qualitätssicherung der Daten, hier speziell der Genotypisierung [5; 6]. Dabei ist die Objektivität dadurch zu gewährleisten, dass die Personen, die die Genotypisierung vornehmen, gegen den Krankheitsstatus verblindet sind. Außerdem wird in der Literatur vorgeschlagen [7], die Reproduzierbarkeit zu sichern, wofür aber allgemeine Empfehlungen schwierig

sind. Denn die Qualität und ihre Sicherung hängen wesentlich von den verwendeten Techniken ab, und diese entwickeln sich stetig weiter.

Im vorliegenden Fall war das Laborpersonal gegenüber dem Fall-Kontroll-Status der Frauen verblindet; außerdem wurde durch eine Wiederholung der Genotypisierung mit derselben Technologie an einer Teilstichprobe die Reproduzierbarkeit gezeigt.

Ein Teil von Genotypisierungsfehlern lässt sich dadurch aufdecken, dass er zu einer Abweichung vom Hardy-Weinberg-Gleichgewicht führt, welches einen gesetzmäßigen Zusammenhang zwischen Allel- und Genotypfrequenzen beschreibt [2]. Daher wird empfohlen, zur Qualitätssicherung auf eine Abweichung in der Kontrollgruppe zu testen. Das Hardy-Weinberg-Gleichgewicht kann jedoch auch durch Effekte verletzt sein, die bei einer krankheitsrelevanten Variante zu erwarten sind. Wenn beispielsweise ein Genotyp die Wahrscheinlichkeit zur Reproduktion verändert, findet eine Selektion statt, und das Gleichgewicht ist nicht mehr gegeben, obwohl die Qualität der Genotypisierung nicht in Frage gestellt werden muss. Da allerdings viele statistische Verfahren die Gültigkeit des Gleichgewichts voraussetzen, sollte die Überprüfung prinzipiell durchgeführt werden.

In der Studie wurde auf Abweichung vom Hardy-Weinberg Gleichgewicht in der Kontrollgruppe getestet, und es wurden keine Abweichungen gefunden.

Über diese für genetische Studien spezifischen Qualitätskriterien hinaus können auch Kriterien wie bei klassischen epidemiologischen Studien angelegt werden. Dazu gehören die standardisierte und einheitliche Definition des Krankheitsstatus' sowie die Absicherung gegen falsch positive Ergebnisse durch eine erfolgreiche Validierung in unabhängigen Gruppen oder durch experimentellen Nachweis der funktionellen Relevanz [7].

Die genaue Definition des Krankheitsstatus wird in der Arbeit nicht gegeben, da die Patientinnen über Registereinträge identifiziert wurden. Für die Kontrollpersonen ist nicht angegeben, auf welche Weise der Krankheitsstatus ausgeschlossen wurde. Die Erhebung anderer Variablen, wie zum Beispiel der Verwendung von Hormonersatztherapie, kann nicht beurteilt werden, weil dazu auf andere Arbeiten verwiesen wird. Eine weitere Bestätigung der Ergebnisse steht noch aus, in der Arbeit ist keine Validierung beschrieben.

Insgesamt ist die Glaubwürdigkeit der Studienergebnisse [4] als gut einzuschätzen, denn die Mehrzahl der Kriterien ist erfüllt.

10b.1.2 Wie lauten die Studienergebnisse?

Die Analyse des genetischen Effekts in einer Assoziationsstudie kann auf einem Vergleich entweder der Allelfrequenzen oder der Genotypfrequenzen zwischen den Gruppen beruhen. Der Vergleich der Allelfrequenzen setzt dabei voraus, dass weder in den Fällen noch in den Kontrollen Abweichungen vom Hardy-Weinberg-Gleichgewicht bestehen [8]. Dasselbe gilt für den Fall, dass aus den genotypischen Daten Haplotypen rekonstruiert oder Haplotyp-Frequenzen geschätzt werden sollen [2]. Bei der Auswahl der Effektschätzer muss also berücksichtigt werden, ob Abweichungen vom Hardy-Weinberg-Gleichgewicht vorliegen. Für die Darstellung und Interpretation der Ergebnisse werden dann in der Regel dieselben Kriterien angelegt wie für klassische epidemiologische Fall-Kontroll-Studien. Dabei werden für das Studiendesign angemessene epidemiologische Maßzahlen mit Konfidenzintervallen ausgewählt. Entsprechend sind bei einer Fall-Kontroll-Studie also Chancenverhältnisse (Odds-Ratios), bei einer Kohortenstudie beispielsweise relative Risiken geeignet, um den Effekt der Assoziation zu beschreiben [2].

Für die univariaten und multivariaten Analysen sind sowohl absolute Häufigkeiten als auch Chancenverhältnisse mit Konfidenzintervallen angegeben. Speziell wird für die Untergruppe der Frauen mit Hormonersatztherapie angegeben, dass sich die Chance für Brustkrebs um das 1,4-fache verringert, wenn sie eine spezifische Variante des Gens tragen.

Um die Relevanz des berichteten Effekts beurteilen zu können, sind zusätzlich weitere Aspekte zu beachten. Erstens ist, wie bei anderen Studiendesigns auch, der geschätzte Effekt nur auf solche Gruppen zu generalisieren, die der untersuchten Stichprobe ähnlich sind. Hier ist bei genetischen Assoziationsstudien spezifisch zu berücksichtigen, dass die Fall- und Kontrollgruppen möglicherweise so ausgewählt wurden, dass der genetische Effekt maximiert wird. Dies bedeutet dann, dass der angegebene Effekt den genetischen Zusammenhang in bevölkerungsbezogenen Gruppen überschätzt. Zusätzlich haben Meta-Analysen von Studien zur genetischen Assoziation gezeigt [9], dass es bei der Erstbeschreibung einer Assoziation häufig zu einer Überschätzung des genetischen Effekts kommt. Entsprechend sind Ergebnisse solcher erster Studien anders zu interpretieren als Effekte aus Validierungsstudien.

Da die Fälle und Kontrollen in dieser Studie bevölkerungsbasiert gezogen wurden, ist nicht von einer Maximierung des genetischen Effekts auszugehen.

Der zweite Aspekt ergibt sich aus dem Problem, dass sich viele berichtete genetische Assoziationen in unabhängigen Studien nicht haben validieren [Validierung] lassen und daher davon auszugehen ist, dass häufig falsch positive Ergebnisse publiziert werden [10]. Für die Einschätzung der Wichtigkeit eines beschriebenen Effekts ist daher neben der eigentlichen Effektschätzung auch die Plausibilität der Assoziation relevant. Entsprechend ist zum einen zu begründen, warum von einer Erblichkeit des untersuchten Phänotyps auszugehen ist [11]. Zum anderen

ist die Auswahl des Gens durch die biologische Funktion oder vorangegangene Studien, beispielsweise systematische Genomscreens oder Genexpressions-Studien, zu rechtfertigen. In ähnlicher Weise kann die Auswahl der Marker innerhalb des Gens entweder mit der vermuteten Funktionalität, der Position innerhalb des Gens oder mit der Struktur des Kopplungs-Ungleichgewichts begründet werden [2]. Zur möglichen Abgrenzung gegen zufällige Befunde müssen außerdem alle durchgeführten statistischen Tests, also alle geprüften Phänotypen und Marker, aufgeführt und ggf. der Umgang mit einem multiplen Testproblem diskutiert werden.

In der Arbeit wird die Auswahl des Kandidatengens selbst plausibel dargestellt, und die genotypisierten Varianten sind bereits aus anderen Studien bekannt. Trotz der Durchführung einer Vielzahl von statistischen Tests in verschiedenen Untergruppen wird aber das Problem des multiplen Testens nicht diskutiert; dies ist umso schwerwiegender, als für die berichtete Assoziation selbst kein p-Wert angegeben wird.

10b.1.3 Sind die Ergebnisse nützlich?

Die Frage nach der Anwendbarkeit und Nützlichkeit für den Patienten stellt sich bei ätiologischen Studien nur begrenzt, weil sie primär nicht mit dem Ziel der direkten Verbesserung der Patientenversorgung durchgeführt werden. Vielmehr steht im Vordergrund, mögliche Ursachen einer Erkrankung zu identifizieren, und die Übertragung in die Klinik findet erst nach weiteren, spezifisch darauf ausgerichteten Studien statt.

Im Fall von genetischen Studien liegt es nahe, dass nach dem Nachweis der Assoziation einer genetischen Variante mit einer Erkrankung dies in die Entwicklung eines diagnostischen Tests – also eines genetischen Tests [genetischer Test] – mündet. Hierfür gelten dann prinzipiell dieselben Kriterien

wie für Screening- und Diagnoseinstrumente (s. Kap. 11) mit der Erweiterung, dass von den Ergebnissen eines genetischen Tests nicht nur der eigentliche Patient, sondern immer auch seine Familie betroffen ist. Dies muss beim Abwägen von möglichem Nutzen und Schaden berücksichtigt werden. Das Kriterium der Nützlichkeit, d.h. dass eine frühe Diagnosestellung auch zu einer Verbesserung im Überleben oder in der Lebensqualität führen muss, erhält hier einen besonderen Stellenwert. So könnte ein positives Testergebnis eine frühere und bessere Überwachung und Prävention auch bei Familienmitgliedern einleiten, aber auch zu einer Stigmatisierung und erhöhten Belastung führen. Auf der anderen Seite könnte ein negatives Testergebnis eine Erleichterung vor allem auch für die Kinder des betroffenen Patienten bewirken. Weitere mögliche Konsequenzen auf der individuellen und gesellschaftlichen Ebene sind ausführlich in der Literatur diskutiert (z.B. [12]), und hilfreiche Hinweise für die Indikation eines genetischen Tests bei Krebserkrankungen finden sich beispielsweise in einer Darstellung der American Society of Clinical Oncology [13].

Literatur

- [1] OMIM (TM). Online Mendelian inheritance in man. <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM> (24.11.06)
- [2] Ziegler A, König IR (2006) A statistical approach to genetic epidemiology. Concepts and applications. Wiley, New York
- [3] Böddeker I, Ziegler A, Assoziations- und Kopplungsstudien zur Analyse von Kandidatengenomen. *Dtsch Med Wschr.* (2000), 125, 810–815
- [4] Le Marchand L et al., MTHFR polymorphisms, diet, HRT, and breast cancer risk: the multiethnic cohort study. *Cancer Epidemiol Biomarkers Prev.* (2004), 13, 2071–2077
- [5] Brosseau L et al., Efficacy of the transcutaneous electrical nerve stimulation for the treatment of chronic low back pain: a meta-analysis. *Spine.* (2002), 27, 596–603
- [6] Vitali SH, Randolph AG, Assessing the quality of case-control association studies on the genetic basis of sepsis. *Pediatr Crit Care Med.* (2005), 6, S74–S77
- [7] Dichgans M, Markus HS, Genetic association studies in stroke: methodological issues and proposed standard criteria. *Stroke.* (2005), 36, 2027–2031
- [8] Sasieni PD, From genotypes to genes: doubling the sample size. *Biometrics.* (1997), 53, 1253–1261
- [9] Ioannidis JP et al., Replication validity of genetic association studies. *Nat Genet.* (2001), 29, 306–309
- [10] Colhoun HM, McKeigue PM, Davey SG, Problems of reporting genetic associations with complex outcomes. *Lancet.* (2003), 361, 865–872
- [11] Cooper DN, Nussbaum RL, Krawczak M, Proposed guidelines for papers describing DNA polymorphism-disease associations. *Hum Genet.* (2002), 110, 207–208
- [12] Guttmacher AE, Collins FS, Welcome to the genomic era. *N Engl J Med.* (2003), 349, 996–998
- [13] American Society of Clinical Oncology policy statement update: genetic testing for cancer susceptibility. *J Clin Oncol.* (2003), 21, 2397–2406

11 Kritische Bewertung von Studien zu diagnostischen Tests

Heiner C. Bucher

Diagnose und Therapie sind Kernaufgaben von Klinikern, deshalb sind Kenntnisse zur kritischen Bewertung von Studien zu diagnostischen Tests wichtig. Hierbei gilt es zu beachten, ob ein diagnostischer Test mit einem Referenztest evaluiert wurde und systematische Fehler (wie z.B. Verification-Bias) in der Evaluation eines diagnostischen Tests vermieden wurden.

Kliniker werden regelmäßig mit neuen diagnostischen Tests und dem Dilemma der Testinterpretation von unerwarteten pathologischen Testergebnissen oder grenzwertigen Befunden konfrontiert. Die technologische Entwicklung in der Medizin erfordert vom Kliniker ein Instrumentarium zur kritischen Evaluation von diagnostischen Tests. Ebenso muss er oder sie klare Konzepte haben, wann Tests für Screening-Untersuchungen und Casefinding geeignet sind. Im Fol-

genden soll anhand eines klinischen Beispiels gezeigt werden, wie eine Studie zu einem diagnostischen Test kritisch evaluiert werden kann. Tabelle 11.1 listet die Kriterien auf, anhand derer vorgegangen werden soll [1].

11.1 Klinisches Szenario

Sie arbeiten als Assistenzärztin auf der Notfallstation einer Großklinik. Es ist ein hektischer Abend und Sie nehmen einen 67-jährigen schwer dyspnotischen Patienten mit bekannter COPD (chronisch obstruktive Lungenkrankheit) auf. Die Anamnese ist schwierig, der Patient klagt über Dyspnoe, die über drei Tage zugenommen habe, sowie Husten mit eitrigem Auswurf. Er verneint Thoraxschmerz. In der klinischen Untersuchung fallen ein spastischer Auskultations-

Tab. 11.1: Kriterien zur Beurteilung der Methodik einer Studie zur diagnostischen Testgenauigkeit [1]

1. Sind die Ergebnisse der Studie glaubwürdig?
<ul style="list-style-type: none"> • Gibt es einen unabhängigen, verblindeten Vergleich mit einem Standardtest (Goldstandard)? • Schloss die Studie ein genügend breites Spektrum von Patienten ein, an denen der Test in der klinischen Praxis angewendet wird? Beeinflussten die Ergebnisse des zu evaluierenden Tests die Entscheidung, ob der Standardtest angewendet wurde? • Wurde die Testmethodik genügend detailliert beschrieben, um die Replikation des Tests zu ermöglichen?
2. Wie lauten die Studienergebnisse?
<ul style="list-style-type: none"> • Werden Wahrscheinlichkeitsverhältnisse der Testergebnisse bzw. Sensitivität und Spezifität angegeben oder die Daten zu deren Berechnung aufgelistet?
3. Sind die Ergebnisse für die Behandlung meiner Patienten nützlich?
<ul style="list-style-type: none"> • Sind die Reproduzierbarkeit und die Interpretation der Testergebnisse in meinem klinischen Umfeld gegeben? • Können die Ergebnisse an meinem Patienten angewendet werden? • Werden die Testergebnisse das Patientenmanagement beeinflussen?

befund mit wenigen niederfrequenten Rasselgeräuschen, leise Herztöne ohne Extratöne, BD von 110/85 mmHg, Puls von 112, O₂-Sättigung von 88%, nicht schlüssig beurteilbarer hepatojugulärer Reflux bei kaltschweißigem Patienten auf. Der Patient ist afebril. Sie geben dem Patienten Sauerstoff und lassen ihn mit Beta-2-Mimetica inhalieren. Das EKG lässt ein akutes koronares Ereignis ausschließen. Nach 30 Minuten geht es dem Patienten etwas besser, die Sauerstoffsättigung beträgt jetzt 92% und Sie erhalten als ersten Laborwert das Brain natriuretic peptide (BNP) mit einem Wert von 250 pg/ml (Normalwert < 50 pg/ml).

Kurz darauf erscheint Ihr Oberarzt und fragt Sie nach der Kurzorientierung, warum Sie aufgrund des erhöhten BNP-Wertes keine Behandlung der Herzinsuffizienz eingeleitet haben. Mit diesem Laborwert sei eine solche Diagnose wahrscheinlich. Er verordnet eine Therapie mit Furosemid (i.v.), Depotnitrat und Morphium 5 mg i.v. Das Zustandsbild des Patienten bessert sich in den darauffolgenden 60 Minuten, was die Verdachtsdiagnose des Oberarztes zu bestätigen scheint. Das Thoraxbild zeigt neben den Zeichen eines Emphysems, einer fraglichen basoapikalen Umverteilung und sehr prominenten Hili keine Auffälligkeiten. Der Patient wird auf die Station verlegt.

Einige Tage später erkundigen Sie sich bei Ihrem Kollegen über den weiteren Verlauf. Sie erfahren, dass die weiteren Abklärungen keine linksventrikuläre Herzinsuffizienz, jedoch eine ausgeprägte pulmonale Hypertonie bei Emphysem ergeben hatten. Eine Exacerbation der COPD sei die wahrscheinliche Diagnose. Sie sind verunsichert und beschließen, Informationen über die diagnostische Wertigkeit des BNP's einzuholen.

11.2 Die Evidenz suchen

Sie suchen mit dem Thesaurus in PubMed und geben die Suchbegriffe (Medical subject headings, so genannte MeSH) „Dyspnoe/

Etiology“ und „Natriuretic Peptide, Brain“ [MeSH] ein. Sie verknüpfen diese Suche mit dem Suchbegriff „Sensitivity and Specificity“ [MeSH] ein und erhalten 71 Treffer. Sie schränken die Suche zusätzlich mit dem Suchbegriff „Prospective Studies“ [MeSH] ein und erhalten 19 Studien. Sie wählen die Arbeit „Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure“ von Maisel et al. aus dem New England Journal of Medicine 2002 [2] aus, da die Studie explizit bei dyspnoischen Patienten, welche sich auf Notfallstationen präsentierten, durchgeführt wurde.

11.3 Die Evidenz bewerten – Kritische Bewertung einer Studie zur diagnostischen Testgenauigkeit

11.3.1 Sind die Ergebnisse der Studie glaubwürdig?

Gibt es einen unabhängigen, verblindeten Vergleich mit einem Standardtest (Goldstandard)?

Ob die Ergebnisse einer Studie vertrauenswürdig sind, hängt von der Studienmethodik und der Testvalidität ab. Es muss somit beurteilt werden, inwiefern der Test und die Testergebnisse die „wahren“ Verhältnisse widerspiegeln. Man sollte sich deshalb versichern, dass ein zu evaluierender Test mit einem angemessenen Referenztest oder -standard verglichen wird. Der Referenztest muss ein Standard sein („Goldstandard“), d.h. der Test mit der höchsten ausgewiesenen Zuverlässigkeit, wie z.B. Biopsie, Autopsie oder Langzeit-Follow-up. Wurde ein Referenztest gewählt, der unbefriedigend ist, dann sind die Studienergebnisse mit hoher Wahrscheinlichkeit nicht brauchbar. In einer Studie wurden beispielsweise D-Dimere als Screening-Test zur Diagnose von tiefen Venenthrombosen des Unterschenkels mit der Duplexsonographie ver-

glichen [3]. Die Verlässlichkeit der Duplexsonographie ist jedoch im Vergleich zur Phlebographie bei der Diagnose der Unterschenkelthrombose geringer, somit ist die Aussagekraft dieser Studie beschränkt.

Falls der Referenztest sinnvoll erscheint, muss weiter überprüft werden, ob die Ergebnisse des Tests und der Referenztest unabhängig evaluiert wurden. D.h. Personen, welche z.B. die Ergebnisse des Tests ablesen, müssen gegenüber den Ergebnissen des Goldstandards (und umgekehrt) verblindet sein. Je größer die Wahrscheinlichkeit ist, dass die Interpretation des Tests durch Kenntnis der Ergebnisse der Referenzuntersuchung beeinflusst werden kann, desto wichtiger ist die Verblindung.

In unserem Beispiel haben die Autoren das Ergebnis des BNP-Tests anhand der Abschlussdiagnose bei allen Patienten, welche sich aufgrund von Dyspnoe auf den Notfallstationen der Studienzentren präsentierten, verglichen. Den Aufnahmeärzten der Notfallstationen war hierbei das Ergebnis des BNP-Tests nicht bekannt. Ob eine Herzinsuffizienz vorlag, wurde anhand aller vorhandenen medizinischen Akten von zwei unabhängigen Kardiologen, die gegenüber den BNP-Werten und den Diagnosen der Aufnahmeärzte verblindet waren, festgestellt. Sie benutzten hierzu die Framingham heart study und den NHANES-Herzinsuffizienz-Score. Bei allen Patienten lagen mindestens zwei bildgebende Verfahren vor: Röntgenthoraxbild und Echokardiographie oder Radionuklid-Ventrikulographie oder Koronar-Angiographie. Eine verbindliche Definition von Test oder Testkombination, welche als Goldstandard galten, fehlte jedoch in der Studie.

Schluss die Studie ein genügend breites Spektrum von Patienten ein, an welchen der Test in der klinischen Praxis angewendet wird?

Ein diagnostischer Test ist nur in dem Ausmaß nützlich, wie er eine gesuchte Ziel-

krankheit bestätigen oder ausschließen lässt, die mit anderen Untersuchungsmethoden nicht differenzierbar ist. Praktisch jeder diagnostische Test kann zwischen Schwerkranken und Gesunden unterscheiden. Der wahre diagnostische Wert einer Untersuchung wird deshalb nur in einer Studie erhellt, welche Patienten einschließt, die die zu suchende Zielkrankheit in unterschiedlicher Ausprägung haben. Dies entspricht ebenso unserer Erfahrung, denn im klinischen Alltag präsentieren sich Krankheitsbilder selten in der typischen Lehrbuchform. Beispielsweise fanden sich bei Patienten mit kolorektalem Karzinom wesentlich höhere Serumspiegel des carcinoembryonalen Antigens (CEA) als bei Normalpersonen. Diese Ergebnisse ließen vermuten, der Test eigne sich als Screening-Test für das kolorektale Karzinom. Folgeuntersuchungen bei einem breiteren Patientenkollektiv mit Patienten mit weniger fortgeschrittenem kolorektalem Karzinom und bei Patienten mit anderen Tumoren oder Grundkrankheiten zeigten hingegen ebenfalls erhöhte CEA-Werte, obwohl gar kein kolorektales Karzinom vorlag (falsch positive Ergebnisse). Der CEA-Test wurde deshalb als Screening-Test für kolorektales Karzinom verworfen.

In unserer Studie wurden 1 586 Patienten mit Dyspnoe aus sieben Zentren eingeschlossen (fünf in den USA und je eines in Frankreich und Norwegen). Patienten mit Dyspnoe, bei denen eine Herzinsuffizienz offensichtlich nicht vorlag (Trauma, Herztamponade), wurden ausgeschlossen. Es ist unklar, ob Patienten konsekutiv in die Studie rekrutiert wurden. Eine Herzinsuffizienz als Schlussdiagnose lag bei 48,5% der Patienten vor. Die Prävalenz der Herzinsuffizienz war somit in dieser Population relativ hoch, dürfte jedoch für eine Patientenpopulation mit akuter Dyspnoe und einem Durchschnittsalter von 64 Jahren repräsentativ sein.

Beeinflussten die Ergebnisse des zu evaluierenden Tests die Entscheidung, ob der Standardtest angewendet wurde?

Die Eigenschaften eines zu evaluierenden Tests werden verzerrt, falls dessen Ergebnisse die Entscheidung beeinflussen, ob Patienten mit dem Referenztest zusätzlich untersucht werden oder nicht. Diese Fehlermöglichkeit wird in der Literatur als „Verification-Bias“ bezeichnet. Ein Beispiel soll dies verdeutlichen. In einer Studie [4] zur Evaluation der Ventilations-Perfusions-Szintigraphie zur Diagnose der akuten Lungenembolie erhielten Patienten mit niedriger Wahrscheinlichkeit einer Lungenembolie im Szintigraphiebefund weniger häufig eine Angiographie (69%), als Patienten mit hochverdächtigem Szintigraphiebefund (92%). Es ist verständlich, dass Kliniker wenig geneigt sind, Patienten weiteren invasiven Untersuchungen auszusetzen, wenn die Chance eines zusätzlichen Informationsgewinn gering ist. Allerdings besteht dabei die Gefahr, dass Patienten mit Lungenembolie trotz negativem Szintigramm falsch zugeordnet werden und die Qualitäten des neuen Tests möglicherweise besser eingeschätzt werden als sie eigentlich sind. Die Autoren dieser Studie korrigierten jedoch diesen Bias, indem sie Patienten mit normalen Szintigraphiebefunden über ein Jahr bezüglich des Auftretens von Lungenembolien nachkontrollierten.

In unserem Beispiel wurden alle eingeschlossenen Patienten, bei denen ein BNP-Wert bestimmt wurde, in die Schlussevaluation durch das unabhängige Kardiologenteam, das den Referenzstandard bildete, einbezogen. Ein Verification-Bias ist somit ausgeschlossen.

Wurde die Testmethodik genügend detailliert beschrieben, um die Replikation des Tests zu ermöglichen?

Studien zu diagnostischen Tests sollten die Testmethoden exakt beschreiben, um deren Replizierbarkeit zu ermöglichen und um

überprüfen zu können, ob der Test unter klinischen Alltagsbedingungen anwendbar ist. Wichtige Angaben zu Technik, Art der verwendeten Geräte oder Labormethoden, Patientenvorbereitung (Diät, Medikation unter der Testsituation etc.) sollten erwähnt sein. Dies ist insbesondere wichtig, wenn der Test stark untersucherabhängig ist.

In unserem Beispiel erhielten alle Patienten eine BNP-Messung. Es wurde ein Immunfluoreszenz-Immunoassay verwendet, und die Autoren beschreiben in einer früheren Publikation die Testanordnung.

11.3.2 Wie lauten die Studienergebnisse?

Werden für die Testergebnisse Wahrscheinlichkeitsverhältnisse angegeben oder die Daten zu deren Berechnung aufgelistet?

Wenn wir einen diagnostischen Test anwenden, gehen wir immer von einem klinischen Verdachtsmoment aus, dass die gesuchte Zielkrankheit vorliegt. Dieser klinische Verdachtsmoment entspricht der Vortest-Wahrscheinlichkeit. Mit dem diagnostischen Test versuchen wir, die Wahrscheinlichkeit für das Vorliegen der Erkrankung zu erhärten oder abzuschwächen. Idealerweise verwenden wir einen Test mit Qualitäten, der uns mit hoher Wahrscheinlichkeit erlauben soll, eine Zielkrankheit möglichst sicher zu bestätigen oder auszuschließen, oder anders formuliert, der eine hohe oder niedrige Nachtest-Wahrscheinlichkeit für das Vorliegen einer Zielkrankheit erzeugt. Die Testqualität hängt von der Sensitivität und Spezifität eines Tests ab. Ein modernes und nützliches Konzept zur Beurteilung von Testqualitäten sind Likelihood-Ratios oder Wahrscheinlichkeitsverhältnisse. Wahrscheinlichkeitsverhältnisse fassen die Testausbeute, die von Sensitivität und Spezifität abhängt, in einer Zahl zusammen.

Das Wahrscheinlichkeitsverhältnis beschreibt die Sicht des Arztes, der ein Tester-

Tab. 11.2: Vergleich der Ergebnisse eines diagnostischen Tests mit einem Standardtest

Ergebnis des Tests unter Evaluation	Referenztest (Standardtest)		
	Krankheit vorhanden	Krankheit nicht vorhanden	
Test positiv	Richtig Positive a	Falsch Positive b	a+b
Test negativ	Falsch Negative c	Richtig Negative d	c+d
	a+c	b+d	a+b+c+d

Sensitivität = $a/(a+c)$ Richtig positive Rate

Spezifität = $d/(b+d)$ Richtig negative Rate

Wahrscheinlichkeitsverhältnis (Likelihood-Ratio) für ein positives Testergebnis = $[a/(a+c)]/[b/(b+d)] = \text{Sensitivität}/(1-\text{Spezifität})$ Richtig positive Rate/Falsch positive Rate

Wahrscheinlichkeitsverhältnis (Likelihood-Ratio) für ein negatives Testergebnis = $[c/(a+c)]/[d/(b+d)] = (1-\text{Sensitivität})/\text{Spezifität}$ Falsch negative Rate/Richtig negative Rate

Positive prädiktiver Wert = $a/(a+b)$ Rate der richtig positiven Testergebnisse

Negativer prädiktiver Wert = $d/(c+d)$

Prävalenz = $a+c/(a+b+c+d)$

Vortest Odds = Prävalenz/(1-Prävalenz)

Nachtest Odds = Vortest Odds x Wahrscheinlichkeitsverhältnis

Nachtest-Wahrscheinlichkeit = Nachtest Odds/(Nachtest Odds +1)

ergebnis zurückbekommt und sich fragt: Wie groß ist die Wahrscheinlichkeit, dass bei diesem (positiven) Testergebnis die Krankheit vorliegt, verglichen mit der Wahrscheinlichkeit, dass trotz des **positiven** Testergebnisses die Krankheit nicht vorliegt? Das Wahrscheinlichkeitsverhältnis für ein positives Testergebnis ist das Verhältnis der richtig positiven Testergebnisse zu den falsch positiven Testergebnissen (vgl. Tab. 11.2). Das Verhältnis der richtig positiven Testergebnisse gemessen an allen Patienten mit der gesuchten

Zielkrankheit entspricht der Sensitivität eines Test ($a/a+c$). Das Verhältnis der falsch positiven Testergebnisse ($b/b+d$) entspricht dem Wert (1-Spezifität). Das Wahrscheinlichkeitsverhältnis für ein negatives Testergebnis ist definiert als das Verhältnis der falsch negativen Testergebnisse zu den richtig negativen Testergebnissen (vgl. Tab. 11.2). Das Verhältnis der falsch negativen Testergebnisse ($c/a+c$) entspricht dem Wert (1-Sensitivität). Das Verhältnis der richtig negativen Testergebnisse, gemessen an allen Patienten, die

Tab. 11.3: Sensitivität, Spezifität, positiver und negativer prädiktiver Wert sowie Wahrscheinlichkeitsverhältnisse (WV) für unterschiedliche Grenzwerte des BNP-Wertes zur Diagnose einer Herzinsuffizienz (Referenz 7)

BNP pg/ml	Sensitivität %	Spezifität %	Positiver prädiktiver Wert %	Negativer prädiktiver Wert %	WV für positives Testergebnis	WV für negatives Testergebnis
50	97	62	71	96	2.5	0.05
80	93	74	77	92	4.0	0.09
100	90	76	79	89	4.3	0.13
125	87	79	80	87	4.35	0.16
150	85	85	83	86	5	0.18

Berechnungsbeispiel des Wahrscheinlichkeitsverhältnisses für einen Grenzwert des BNP von 100 pg/ml:

$WV + = \text{Sensitivität} / (1 - \text{Spezifität}) = 90\% / (100\% - 76\%) = 4.3$

$WV - = (1 - \text{Sensitivität}) / \text{Spezifität} = (100\% - 90\%) / 76\% = 0.13$

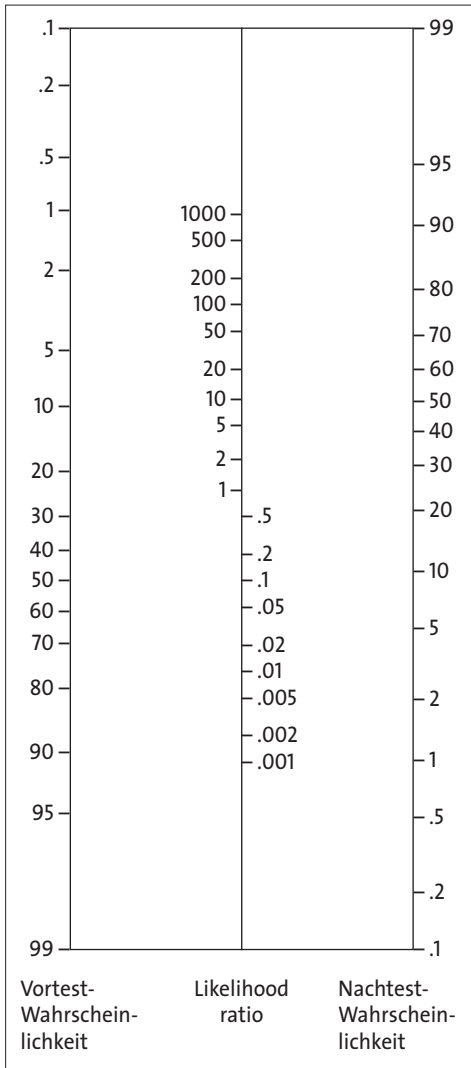


Abb. 11.1: Likelihood-Ratio (nach [5]) zur Berechnung von Nachtest-Wahrscheinlichkeiten

nicht erkrankt sind, entspricht der Spezifität eines Tests ($d/b+d$).

Der große Vorteil des Wahrscheinlichkeitsverhältnisses ist, dass mittels der Vortest-Wahrscheinlichkeit unmittelbar die Nachtest-Wahrscheinlichkeit eines Testergebnisses berechnet werden kann. Die Vortest-Wahrscheinlichkeit wird auch als die Prävalenz einer Zielkrankheit bezeichnet.

Kehren wir zu unserem Beispiel zurück. In Tabelle 11.3 sind die Sensitivität, Spezifi-

tät und der positive und negative prädiktive Wert für verschiedene angenommene Grenzwerte des BNP aus der Originalpublikation aufgeführt. Bei einem niedrigeren angenommenen Grenzwert wird die Sensitivität respektive der negative prädiktive Wert des BNP-Tests optimiert, während bei einem höheren angenommenen Grenzwert des BNP-Tests die Spezifität und der positive prädiktive Wert optimiert werden. Bei einem Grenzwert von 100 pg/ml ergibt sich, wie die einfachen Berechnungen in Tabelle 11.3 zeigen, ein Wahrscheinlichkeitsverhältnis von 4.3 für einen positiven, respektive von 0.13 für einen negativen Test. Ein Wahrscheinlichkeitsverhältnis von 4.3 besagt, dass eine Herzinsuffizienz bei einem Grenzwert von 100 pg/ml mit einer 4.3-mal größeren Wahrscheinlichkeit bei einem Patienten vorkommt, bei dem die diagnostische Abklärung tatsächlich eine Herzinsuffizienz zeigt als bei einem Patienten, der in der diagnostischen Abklärung keine Herzinsuffizienz hat. Entsprechend besagt ein Wahrscheinlichkeitsverhältnis von 0.13, dass ein solches Ergebnis viel wahrscheinlicher bei einem Patienten zu erwarten ist, bei dem sich in der diagnostischen Abklärung keine Herzinsuffizienz nachweisen lässt. Je kleiner das Verhältnis, desto höher ist die Wahrscheinlichkeit, dass bei einem negativen Testergebnis die Krankheit auch tatsächlich nicht vorliegt. Anhand des Nomogramms von Fagan [5] können die Nachtest-Wahrscheinlichkeiten leicht abgelesen werden (s. Abb. 11.1).

Wir suchen die Vortest-Wahrscheinlichkeit auf der linken Skala auf (z.B. die Schätzung von 50%), ziehen eine Linie zur mittleren Skala mit den Wahrscheinlichkeitsverhältnissen (in unserem Beispiel 4.3 für einen positiven Test und 0.13 für einen negativen BNP-Test bei einem Grenzwert von 100 pg/ml) und lesen auf der rechten Skala die Nachtest-Wahrscheinlichkeit ab. In unserem Beispiel beträgt die Nachtest-Wahrscheinlichkeit für einen positiven Test 81% und für

Online Clinical Calculator

Bayesian Analysis Model

Enter PREVALENCE, SENSITIVITY, and SPECIFICITY:

PREV: SENS: SPEC:

Or enter TP, FN, TN, and FP.

	Disease	No Disease
Positive Test	TP: <input type="text"/>	FP: <input type="text"/>
Negative Test	FN: <input type="text"/>	TN: <input type="text"/>

(+) PREDICTIVE VALUE =

(-) PREDICTIVE VALUE =

(+) LIKELIHOOD RATIO =

(-) LIKELIHOOD RATIO =

Remember:

$Sensitivity = TP / (TP + FN)$

Abb. 11.2: Likelihood-Ratio-Kalkulator [<http://www.intmed.mcw.edu/clincalc/bayes.html>]

einen negativen Test 13,8%. Verschiedene Programme wie *epocrates.com*, welches auf einen Palm geladen werden kann und rasch am Krankenbett oder in der Ambulanz benutzt werden kann, erlauben die einfache und rasche Berechnung von Nachttest-Wahrscheinlichkeiten mittels der Werte von Sensitivität und Spezifität. Im Internet gibt es ebenfalls verschiedene frei zugängliche Kalkulatoren (z.B. <http://www.intmed.mcw.edu/clincalc/bayes.html> – s. Abb. 11.2).

Unser Beispiel macht deutlich, wie die Testausbeute von der Vortest-Wahrscheinlichkeit abhängt. Die prädiktive Wertigkeit eines Tests hängt entscheidend von der Vortest-Wahrscheinlichkeit ab. Je höher bei gegebener Sensitivität und Spezifität die Vortest-Wahrscheinlichkeit, desto höher die prädiktive Wertigkeit des Tests. Die Nachttest-Wahrscheinlichkeit entspricht dem positiven respektive negativen prädiktiven Wert eines Tests. Der positive prädiktive Wert ei-

nes Tests entspricht dem Verhältnis der richtig positiven Testergebnisse an der Gesamtzahl aller positiven Testergebnisse (vgl. Tab. 11.2: $a/a+b$). Dementsprechend entspricht der negative prädiktive Wert eines Tests dem Verhältnis der richtig negativen Testergebnisse an der Gesamtzahl aller negativen Testergebnisse (vgl. Tab. 11.2: $d/c+d$). Allerdings ist die Spezifität des BNP eher mittelmäßig, weswegen der positive prädiktive Wert des BNP in Abhängigkeit des Grenzwertes deutlich schlechter ist als der negative prädiktive Wert.

Ein Nachteil von Wahrscheinlichkeitsverhältnissen ist, dass die Berechnung der Nachttest-Wahrscheinlichkeiten etwas umständlich und nicht aus der Hand durchzuführen ist. (Die genaue Formel ist in anderen Quellen ausführlicher beschrieben, z.B. [6]. Das Likelihood-Ratio Nomogramm von Fagan (s. Abb. 11.1.) oder ein Kalkulationsprogramm (s. Abb. 11.2.) können hier helfen.

Wahrscheinlichkeitsverhältnisse haben jedoch gegenüber den herkömmlichen Konzepten von Sensitivität und Spezifität mehrere Vorteile. Sie erlauben eine flexiblere Testinterpretation, da erstens die Vortest-Wahrscheinlichkeit mitberücksichtigt werden kann und zweitens die Testeigenschaften und Nachtest-Wahrscheinlichkeiten nicht anhand eines willkürlichen Grenzwertes evaluiert werden müssen, sondern anhand verschiedener gewählter Grenzwerte besser beurteilt werden können. Tabelle 11.3 verdeutlicht, wie bei zunehmend höherem Grenzwert des BNP sowohl das Wahrscheinlichkeitsverhältnis als auch der positive prädiktive Wert des Tests zunehmen. Umgekehrt zeigt sich, dass bei niedrigen BNP-Grenzwerten das Wahrscheinlichkeitsverhältnis für einen negativen Test abnimmt und der negative prädiktive Wert des Tests deutlich besser wird.

Tests mit großen Wahrscheinlichkeitsverhältnissen erhöhen – bei gegebener Vortest-Wahrscheinlichkeit – die Wahrscheinlichkeit, dass die gesuchte Krankheit vorhanden ist. Umgekehrt machen Tests mit einem niedrigen Wahrscheinlichkeitsverhältnis das Vorliegen einer gesuchten Krankheit unwahrscheinlich. Als Faustregel gilt, dass Wahrscheinlichkeitsverhältnisse in der Größenordnung von >10 bei positivem Test oder <0.1 bei negativem Test zu Nachtest-Wahrscheinlichkeiten führen, welche die Vortest-Wahrscheinlichkeit deutlich verändern und eine gesuchte Zielkrankheit oft sicher ein-, respektive ausschließen lassen. Wahrscheinlichkeitsverhältnisse zwischen 5 und 10 respektive 0.1 und 0.2 ergeben in der Regel nur mäßige Veränderungen der Nachtest-Wahrscheinlichkeit. Wahrscheinlichkeitsverhältnisse zwischen 2 und 5 respektive 0.5 und 0.2 bewirken geringe, jedoch manchmal wichtige Veränderungen der Nachtest-Wahrscheinlichkeit. Wahrscheinlichkeitsverhältnisse zwischen 1 und 2 und 0.5 und 1 verändern die Nachtest-Wahr-

scheinlichkeit in einem klinisch nicht relevanten Ausmaß.

Unser Beispiel zeigt, dass die Wahrscheinlichkeitsverhältnisse des BNP ungenügend sind, um die Diagnose einer Herzinsuffizienz mit großer Sicherheit zu bestätigen. Hingegen ist bei einem Grenzwert von 50 pg/ml die Sensitivität (97%) des BNP sehr gut. Bei einem Test mit einer hohen Sensitivität lässt sich bei einem negativen Testergebnis eine gesuchte Zielkrankheit mit großer Wahrscheinlichkeit ausschließen. Eselsbrücke: SnNout (engl. Schnauze): Sensitivity high: Negative test result rules out the target disorder. Umgekehrt lässt sich bei einem Test mit einer hohen Spezifität, also bei einem positiven Testbefund, die Zielkrankheit mit großer Wahrscheinlichkeit einschließen. Eselsbrücke: SpPin: Specificity high: Positive test result rules in the target disorder.

Bei Vortest-Wahrscheinlichkeiten von 30%–50% und einem negativen BNP-Test (<50 pg/ml; WV: $(100\% - 97\%) / 62\% = 0.05$) resultieren sehr niedrige Nachtest-Wahrscheinlichkeiten, welche das Vorliegen einer Herzinsuffizienz mit hoher Wahrscheinlichkeit ausschließen lassen.

Die Autoren haben zusätzlich überprüft, ob sich durch den Einbezug von weiteren klinischen Symptomen die positive prädiktive Wertigkeit des Tests verbessern ließe. Hierbei zeigt sich in einem multivariaten Modell, dass das Vorliegen einer positiven Anamnese für eine Herzinsuffizienz (Odds-Ratio 11.1) und eine baso-apicale Umverteilung im Thoraxröntgenbild (Odds-Ratio 10.7) nebst einem BNP >100 pg/ml (Odds-Ratio 29.6), die wichtigsten prädiktiven Faktoren für das Vorliegen einer Herzinsuffizienz waren.

Vereinfachend können wir auch sagen, dass die Testausbeute von zwei Faktoren bestimmt wird: dem Wahrscheinlichkeitsverhältnis und der Vortest-Wahrscheinlichkeit. Je höher respektive niedriger das Wahrscheinlichkeitsverhältnis, desto besser ist die Testausbeute. Der zweite entscheidende Fak-

tor ist die Vortest-Wahrscheinlichkeit: Bei einer Vortest-Wahrscheinlichkeit von 40%–60% und gegebener Sensitivität und Spezifität ist die Testausbeute am höchsten, d.h. es werden die besten Nachtest-Wahrscheinlichkeiten erzielt, die eine gesuchte Zielkrankheit sicher ausschließen respektive bestätigen lassen.

Ein weiterer Vorteil von Wahrscheinlichkeitsverhältnissen ist, dass die Ausbeute bei Anordnung von sequenziellen Tests analysiert werden kann. Hier geht es um Fragen wie z.B.: Was bringt eine Radionuklidventrikulographie an Zusatzinformation zu einem grenzwertigen Befund im Belastungs-EKG? In diesem Fall entspricht die Nachtest-Wahrscheinlichkeit des Belastungs-EKGs der Vortest-Wahrscheinlichkeit der Radionuklid-Ventrikulographie. Die Bedeutung des Wahrscheinlichkeitsverhältnisses zur Optimierung der klinischen Entscheidungsfindung ist in Zunahme begriffen. Es besteht bereits eine ausgezeichnete Sammlung zu Testevaluationen mit Likelihood-Ratios für häufige klinische Probleme [7]. Zudem wird die Cochrane Collaboration in Kürze systematische Übersichtsarbeiten zu diagnostischen Tests in die Cochrane Library aufnehmen.

11.3.3 Sind die Ergebnisse für die Behandlung meiner Patienten nützlich?

Sind die Reproduzierbarkeit und die Interpretation der Testergebnisse in meinem klinischen Umfeld gegeben?

Eine wichtige Eigenschaft eines diagnostischen Tests ist dessen Reproduzierbarkeit, d.h. die Eigenschaft, z.B. bei stabilen Patienten oder Testkonditionen die gleichen Ergebnisse zu liefern. Eine schlechte Reproduzierbarkeit kann mit technischen oder methodischen Problemen eines Tests (z.B. Radioimmuno-Assay) zusammenhängen. Eine unterschiedliche Testinterpretation von Untersuchern kann

ebenso die Reproduzierbarkeit eines Tests beeinflussen. Deshalb sollten Angaben zur Reproduzierbarkeit vorhanden sein. Dies ist besonders dann wichtig, wenn die Testinterpretation spezielle Fähigkeiten der Untersucher erfordert (z.B. Interpretation von Computertomographien, Ultraschalluntersuchungen, Elektrokardiogramm etc.). Wenn die Reproduzierbarkeit eines Tests nur mittelmäßig ist und der Test dennoch zwischen pathologischen und normalen Befunden unterscheiden hilft, besteht weniger Anlass zur Sorge. Ist die Reproduzierbarkeit eines Test sehr hoch und die Beurteilung durch die Untersucher variiert nur geringfügig, dann ist der Test in seiner Anwendung entweder einfach oder die Testinterpreten sind hoch qualifiziert. Falls letzteres zutrifft, kann der Test in einem anderen klinischen Umfeld eine geringere Reproduzierbarkeit haben.

Angaben zur Reproduzierbarkeit fehlen in unserem Beispiel gänzlich. Im konkreten Fall muss gegebenenfalls mittels Nachfrage vor Ort überprüft werden, ob Qualitätsstandards oder Daten vorhanden sind, welche eine Beurteilung der Reproduzierbarkeit von diagnostischen Tests in unserem eigenen klinischen Umfeld zulassen.

Können die Ergebnisse an meinem Patienten angewendet werden?

Testeigenschaften können sich ändern, wenn diagnostische Tests in unterschiedlichen Patientenpopulationen angewendet werden, die sich z.B. im Schweregrad der gesuchten Zielkrankheit unterscheiden. Ebenso können Störfaktoren wie z.B. Komorbidität, die das Testergebnis beeinflussen, variieren. Wenn die zu untersuchende Patientenpopulation kränker ist, d.h. eine ausgeprägtere Form der gesuchten Zielkrankheit hat, nimmt die Testsensitivität zu, was bedeutet, dass der numerische Wert des Wahrscheinlichkeitsverhältnisses steigt und sich weiter von 1 entfernt. Falls die Patienten nur leichte Affektionen der Zielkrankheit aufweisen, be-

wegt sich das Wahrscheinlichkeitsverhältnis näher zu 1 (die Sensitivität nimmt ab). Falls Patienten ohne die gesuchte Zielkrankheit Faktoren aufweisen, die ein falsch-positives Ergebnis begünstigen, dann bewegt sich das Wahrscheinlichkeitsverhältnis ebenfalls näher zu 1, und die Aussagekraft des Tests nimmt wiederum ab. Falls umgekehrt Faktoren, die das Testergebnis verfälschen können, in der Zielpopulation weniger häufig sind, nimmt das Wahrscheinlichkeitsverhältnis zu und die Aussagekraft des Tests steigt. Dass sich Testeigenschaften in der Zielpopulation, die unterschiedliche Erkrankungsausprägungen aufweisen, ändern, ist gut dokumentiert. Beispielsweise fanden sich bei Patienten mit ausgeprägteren Koronarskiosen bei positiven Belastungs-EKGs höhere Wahrscheinlichkeitsverhältnisse [8].

Falls die Verhältnisse zur Testdurchführung ähnlich sind wie in der Studie und Ihr Patient unter die gleichen Ein- und Ausschlusskriterien fällt wie die Studienpopulation, dann sind die Studienergebnisse auf die eigene Patientenpopulation anwendbar. Wenn dies nicht zutrifft, dann muss aufgrund der Testeigenschaften und der Situation in Klinik oder Praxis entschieden werden, inwiefern der jeweilige diagnostische Test in Ihrer klinischen Praxis tauglich ist.

Die Studie lässt nur begrenzt eine Aussage zu, ob es sich bei der eingeschlossenen Studienpopulation um eine hoch selektionierte Population handelt. Die Prävalenz der Herzinsuffizienz war mit 48,5% relativ hoch, ein Drittel der Patienten hatte eine anamnestisch bekannte Herzinsuffizienz, ein Viertel hatte einen Myokardinfarkt, 40% eine COPD. Sie schließen hieraus, dass Ihr Patient mit der hier untersuchten Studienpopulation vergleichbar ist, und dass die Testcharakteristika etwa ähnlich sein könnten wie bei der Studienpopulation.

Werden die Testergebnisse das Patientenmanagement beeinflussen?

Unsere Ausführungen sollen zeigen, inwiefern explizite Überlegungen zu Wahrscheinlichkeiten des Vorhandenseins einer gesuchten Zielkrankheit die klinische Entscheidungsfindung bei diagnostischen Problemen leiten können. Es ist wichtig, dass für jede Zielkrankheit, die der Test identifizieren soll, Schwellenwerte definiert werden, deren Unter- beziehungsweise Überschreitung keine weitere Abklärung beziehungsweise eine Weiterabklärung oder direkte Behandlung nach sich zieht.

11.4 Auflösung des klinischen Szenarios

Ausgerüstet mit dieser Information machen Sie sich an die Lösung Ihres klinischen Problems. Nehmen wir an, Sie hätten bei der initialen Beurteilung des Patienten das Vorliegen einer Herzinsuffizienz mit 30% ($p=0.3$) geschätzt. Beim gewählten Grenzwert von 50 pg/ml und einem Wahrscheinlichkeitsverhältnis von 2.5 ergibt sich eine Nachtest-Wahrscheinlichkeit von rund 51%. Dies ergibt sich aus dem Nomogramm oder aus der Berechnung gemäß der Formeln aus Tabelle 11.2 ($(0.3/0.7) \times 2.5 = 1.07$ und $1.07 / (1.07 + 1) = 0.51$). Trotz des leicht erhöhten BNP-Wertes und eines relativ hohen positiven prädiktiven Wertes bestätigte sich in der weiteren Abklärung bei diesem Patienten die Diagnose einer Linksherzinsuffizienz nicht. Der erhöhte BNP-Wert ist im Zusammenhang mit der später gefundenen pulmonal-arteriellen Hypertonie zu sehen. Das Beispiel illustriert, dass das Wahrscheinlichkeitsverhältnis für einen positiven BNP-Test zu wenig hoch ist, um das Vorhandensein einer Herzinsuffizienz sicher zu diagnostizieren. Die relativ hohe Sensitivität des Tests macht diesen Test jedoch zu einem nützlichen Instrument für die Ausschlussdiagnose einer Herzinsuffizienz. Eine Meta-Analyse von diagnostischen Studien zeigt, dass die gepoolte nega-

tive Wahrscheinlichkeitsrate für den BNP-ELISA-Test 0.12; (95% CI, 0.09–0.16) und für den RIA-Test 0.23 (95% CI, 0.16–0.32) beträgt [9]. Dies unterstreicht, dass insbesondere der BNP-ELISA-Test ein nützlicher Test zum Ausschluss einer Herzinsuffizienz ist. In diesem Sinne können Sie nun den BNP im klinischen Management von Patienten differenzierter einsetzen und die Wertigkeit von BNP-Tests besser interpretieren.

Literatur

- [1] Jaeschke R, Guyatt GH, Sackett DL, Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA. (1994), 271, 703–707
- [2] Maisel AS et al., Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. N Engl J Med. (2002), 347, 161–167
- [3] Harvey RL et al., Deep vein thrombosis in stroke. The use of plasma D-dimer level as a screening test in the rehabilitation setting. Stroke. (1996), 27, 1516–1520
- [4] Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the prospective investigation of pulmonary embolism diagnosis (PIOPED). The PIOPED Investigators. JAMA. (1990), 263, 2753–2759
- [5] Fagan TJ, Letter: Nomogram for Bayes theorem. N Engl J Med. (1975), 293, 257
- [6] Sackett DL et al. (1991) Clinical epidemiology. A basic science for clinical medicine. Lippincott Williams and Wilkins, London
- [7] Black ER et al. (1999) Diagnostic strategies for common medical problems. John Hopkins Univ. Pr., Philadelphia
- [8] Hlatky MA et al., Factors affecting sensitivity and specificity of exercise electrocardiography. Multivariable analysis. Am J Med. (1984), 77, 64–71
- [9] Battaglia M et al., Accuracy of B-type natriuretic peptide tests to exclude congestive heart failure: systematic review of test accuracy studies. Arch Intern Med. (2006), 166, 1073–1080

12 Kritische Bewertung von Studien zu therapeutischen Interventionen

Hilke Bertelsmann, Guido Lerzynski, Regina Kunz

12.1 Klinisches Szenario

In ihrer Allgemeinarztpraxis behandelt Frau Dr. E. Vident seit mittlerweile vier Jahren Herrn Hans Süßlich, einen 57-jährigen Buchhalter, der außer an Übergewicht (BMI 29) seit drei Jahren an einem manifesten Diabetes mellitus Typ 2 leidet und nach anfänglicher oraler antidiabetischer Therapie seit zwei Monaten insulinpflichtig geworden ist. Darüber hinaus wird Herr Süßlichs bereits seit zehn Jahren bekannter Bluthochdruck mit einem Diuretikum und einem ACE-Hemmer behandelt. Dennoch finden sich bei dem in der Sprechstunde kontrollierten Blutdruck regelmäßig Werte um 150/90 mmHg. Die Lipidwerte sind unter Statinbehandlung gut eingestellt (LDL 124 mg/dl).

Vor vier Monaten berichtete Herr Süßlich erstmalig von einer nach Anstrengungen auftretenden Luftnot und Enge in der Brust. Die daraufhin durchgeführten Untersuchungen lieferten Hinweise auf das Vorliegen einer koronaren Herzerkrankung und wurden gründlich diagnostisch abgeklärt. Zurzeit erscheint eine medikamentöse Therapie mit Betablockern, Nitraten und Aspirin ausreichend, um das Risiko für einen Herzinfarkt oder Schlaganfall zu senken.

Herr Süßlich ist Nichtraucher. In seiner Freizeit liest er viel, ist vielseitig interessiert und nutzt auch das Internet, um sich über interessante Themen zu informieren. Zur Kontrolluntersuchung bringt er den Ausdruck einer Patienteninformation mit, die er im Internet gefunden hat. Dort wird eine Therapie mit Aspirin und zusätzlich Clopidogrel empfohlen, die angeblich gerade in großen internationalen Studien getestet werde und sehr vielversprechende Erfolge bei der Verhinderung von Herzinfarkten bringe. Herr Süßlich möchte

wissen, ob er diese zusätzliche Prophylaxe nicht auch einnehmen solle. Ein zusätzliches Medikament erscheint ihm attraktiver als die mühsam eingehaltene Diät und das empfohlene Bewegungsprogramm. Frau Dr. E. Vident hat von dieser Therapie bisher noch nichts gehört, verspricht aber, sich zu informieren. Sie beschließt, dies nach den Methoden der Evidenzbasierten Medizin, die sie kürzlich in einem Einführungskurs kennengelernt hat, durchzuführen.

Im Gespräch mit einem befreundeten Kollegen wird anhand des PICO-Formats (s. Kap. 8) folgende Fragestellung definiert: Gibt es bei Patienten mit Diabetes und koronarer Herzerkrankung (KHK) für die Kombinationstherapie Clopidogrel/Aspirin im Vergleich zur alleinigen Gabe von Aspirin einen zusätzlichen Nutzen bzgl. der Prävention von kardiovaskulären Ereignissen?

12.2 Die Evidenz suchen

Frau Dr. E. Vident und ihr Kollege durchsuchen zuerst einmal die Cochrane Library (<http://www.cochrane.org/reviews/index.htm>) nach systematischen Reviews zu ihrer Fragestellung. Nachdem sie hier nicht fündig wurden, suchen sie in der öffentlich zugänglichen Datenbank Pubmed (MEDLINE; www.ncbi.nlm.nih.gov/entrez) nach Primärstudien. Sie geben die Suchworte „Clopidogrel“, „Aspirin“ in die Freitextsuche ein und erhalten insgesamt 1373 Treffer, die thematisch noch wenig einheitlich erscheinen. Mit der Funktion „Limits“ schränken sie ihre Suche auf Artikel in Deutsch oder Englisch ein. Sie wählen mit der Einschränkung „human“ Studien an Menschen und wählen unter der

Funktion „Type of Article“ „**Randomized Controlled Trial**“ aus. Mit diesen Einschränkungen erbringt die Suche nur noch 125 Artikel. Bei der Durchsicht der ersten Treffer wird klar, dass die gefundenen Studien sich auf unterschiedliche Zielereignisse beziehen. Frau Dr. Vident und ihr Kollege geben zusätzlich den Suchbegriff „**Atherothrombotic Events**“ in die Freitextsuche ein. Unter den 71 Treffern finden sie an dritter Stelle der Liste eine Studie, die ihrer Fragestellung optimal entspricht:

Bhatt DL et al. for the CHARISMA Investigators. **Clopidogrel and aspirin versus aspirin alone for the prevention of atherothrombotic events.** N Engl J Med. 2006 Apr 20; 354 (16): 1706–17. Epub 2006 Mar 12. [1]

In dieser Studie wurde die Wirksamkeit einer Kombinationstherapie aus Aspirin (75–162 mg pro Tag und Clopidogrel 75 mg pro Tag als präventive Maßnahme zur Vermeidung von kardiovaskulären Ereignissen für Patienten mit kardiovaskulären Risikofaktoren untersucht.

Die beiden Kollegen bestellen diese hochaktuelle Studie, die gerade erst im New England Journal of Medicine erschienen ist, und machen sich an die Auswertung. Im weiteren Verlauf sprechen sie von dieser Studie als „CHARISMA-Studie“.

12.3 Die Evidenz bewerten

12.3.1 Sind die Ergebnisse der Studie glaubwürdig?

Eine klinische Studie zur Wirksamkeit einer Therapie oder einer präventiven Maßnahme untersucht den Zusammenhang zwischen der Therapie und einer oder mehreren Zielvariablen (wie z.B. Herzinfarkt, Schlaganfall oder Todesfälle) an einer definierten Patienten-Population.

Frau Dr. Vident möchte wissen, ob Herr Süßlich von einer Kombinationstherapie aus

Aspirin und Clopidogrel profitieren würde. Die vorliegende Studie untersucht den Zusammenhang zwischen dieser Therapie im Vergleich zu einer Therapie mit Aspirin allein und dem Risiko, ein kardiovaskuläres Ereignis (Herzinfarkt, Schlaganfall oder kardiovaskulärer Todesfall) zu bekommen. Dieser „wahre Wert“ der Grundpopulation, der den Unterschied des Risikos für ein kardiovaskuläres Ereignis zwischen einer Gruppe von Patienten mit Kombinationstherapie und Patienten mit Aspirin-Monotherapie symbolisiert, ist in der Realität natürlich nicht bekannt.

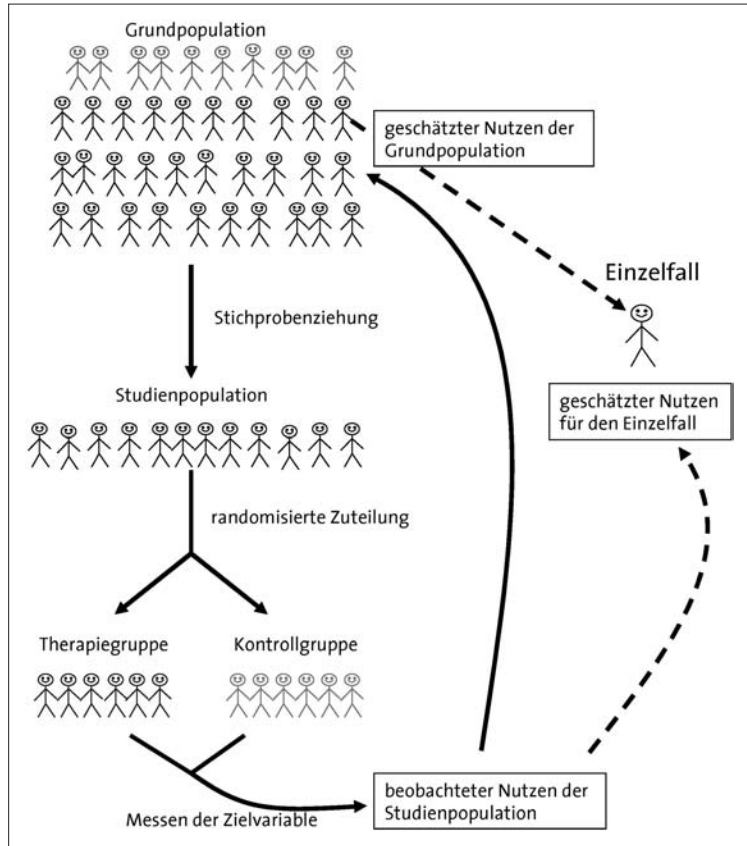
Das an der Studienpopulation gemessene Ergebnis wird deswegen als Schätzwert für diese Grundpopulation verwendet. Je besser die Studie aufgebaut ist, desto sicherer kann von dem gemessenen Ergebnis auf den „wahren Wert“ der Grundpopulation geschlossen werden. Frau Dr. Vident wird dann den aus der Studie geschätzten Risikounterschied der Grundpopulation auf das individuelle Risiko ihres Patienten übertragen. Je eher man Herrn Süßlich als einen typischen Vertreter der Grundpopulation betrachten kann, desto besser wird die Übertragung seinem individuellen Risiko entsprechen.

Um diese Übertragung möglichst sicher und zum Nutzen des Patienten durchführen zu können, muss die Studie in mehreren Schritten bewertet werden. Im Folgenden wird zuerst untersucht, ob die Ergebnisse mit hoher Wahrscheinlichkeit den Nutzen der Grundpopulation widerspiegeln. Im zweiten Schritt geht es um die Frage, was die Ergebnisse über den Nutzen der Therapie selbst aussagen.

Das in einer Studie beobachtete Ergebnis kann grundsätzlich aus zwei Gründen vom „wahren Wert“ der Grundpopulation abweichen:

- ▲ durch Zufallsfehler
 - Zufallsfehler entstehen, weil die Personen und Messwerte in der Stichprobe zufällig vom Durchschnitt der Grund-

Abb. 12.1: Übertragung des Studienergebnisses auf die Grundpopulation und den Einzelfall



population abweichen. Angaben über die Power der Studie und die statistische Signifikanz der Ergebnisse (siehe unten) helfen, die Wahrscheinlichkeit für einen Zufallsfehler einzuschätzen.

- ▲ durch systematische Fehler (Bias)
 - Systematische Fehler entstehen, wenn sich das Risiko, ein Zielereignis wie z.B. einen Herzinfarkt oder Schlaganfall zu erleiden, in der Therapiegruppe und der Kontrollgruppe zu Beginn oder im Verlauf der Studie unterscheidet, ohne dass dieser Unterschied durch die Therapie bedingt ist. Die Wahrscheinlichkeit für einen systematischen Fehler kann anhand der methodischen Qualität des Studiendesigns eingeschätzt werden.

Die Frage nach der Glaubwürdigkeit der Studie ist erst einmal die Frage nach der Wahrscheinlichkeit für einen systematischen Fehler. Auf die Wahrscheinlichkeit für einen Zufallsfehler wird bei der Betrachtung der Ergebnisse eingegangen. Die Frage nach einem systematischen Fehler wird beantwortet, indem das Studiendesign detailliert betrachtet wird. Die Planung und Durchführung einer Studie sollte sicherstellen, dass Fehlerquellen, die zu Abweichungen des gemessenen Risikounterschiedes von dem „wahren Wert“ der Grundpopulation führen könnten, weitgehend vermieden werden.

Eine erste Einschätzung der Validität einer Studie ist anhand der so genannten Evidenzstufen möglich. Je höher ein Studientyp auf der Evidenzstufe angeordnet ist, desto glaubwürdiger sind die Studienergebnisse.

Tab. 12.1: Hierarchie der Evidenzstufen (<http://cebmr2.ox.ac.uk>)

Evidenzstufe	Studienart
Ia	Systematische Übersicht von randomisierten, kontrollierten Studien (mit Homogenität)
Ib	Einzelne randomisierte, kontrollierte Studie mit engem Konfidenzintervall
IIa	Systematische Übersicht aus Kohortenstudien (mit Homogenität)
IIb	Einzelne Kohortenstudie/randomisierte kontrollierte Studie mit methodischen Mängeln
IIc	Outcome-Forschung
IIIa	Systematische Übersicht aus Fall-Kontroll-Studien
IIIb	Einzelne Fall-Kontroll-Studie
IV	Fall-Serien, Kohortenstudien mit methodischen Mängeln
V	Konsensuskonferenzen und/oder klinische Erfahrungen anerkannter Autoritäten ohne explizite Grundlage von kritisch bewerteter Evidenz

Allerdings kann die Glaubwürdigkeit einer Studie durch zusätzliche Fehlerquellen im Studiendesign eingeschränkt werden, auch wenn der Studientyp grundsätzlich einer hohen Evidenzstufe zugeordnet ist.

Die Stufen der Evidenztreppe veranschaulichen die von unten nach oben zunehmende Aussagekraft und Glaubwürdigkeit der Ergebnisse verschiedener Studientypen. Die Fallserien der Stufe IV stellen im Gegensatz zu den Expertenmeinungen, physiologischen Überlegungen und Tierstudien auf Stufe V systematisch dokumentierte Beobachtungen an betroffenen Patienten dar, die eine bestimmte Therapie erhalten haben. Allerdings bleibt die Frage offen, wie sich der Gesundheitszustand der Patienten ohne die Therapie entwickelt hätte, sodass Verbesserungen nicht sicher auf die Therapie zurückgeführt werden können. Alle Studien ab Stufe III beobachten deswegen parallel zu der Therapiegruppe eine Kontroll-, Placebo- oder Standardtherapiegruppe. Erst der Unterschied in der Zielereignisrate zwischen Therapie- und Kontrollgruppe (in diesem Beispiel die Rate der Herzinfarkte) ermöglicht eine Einschätzung des Therapieeffektes. Die retrospektiven Studien der Gruppe III (Fall-Kontroll-Studien) vergleichen in der Regel eine Gruppe von Patienten, bei denen das

Zielereignis (z.B. der Herzinfarkt) bereits eingetroffen ist, mit einer Gruppe Patienten ohne Zielereignis. Sie ermitteln, wie diese Patienten in der Vergangenheit behandelt wurden und ob ein Behandlungsunterschied in der Vergangenheit das Auftreten von Herzinfarkten in der Kontrollgruppe erklärt. Dieses Studiendesign ist schnell und kostengünstig, birgt aber das Problem, dass die frühere Therapie oft nur lückenhaft aus Dokumenten oder Erinnerungen beteiligter Personen rekonstruiert werden kann (Recall-Bias). Auch die Auswahl der Patienten für die Studie entspricht häufig nicht dem Durchschnitt aller therapierten Patienten (Selektions-Bias). Als besonders schwierig erweist sich die Auswahl der Kontrollen, die alle Personen mit der entsprechenden Erkrankung repräsentieren sollen, die die Therapie nicht erhalten. Durch die Auswahl besonders kranker oder auch besonders interessierter Kontrollpersonen kann ebenfalls ein Selektions-Bias entstehen. Beide Biasformen können zu verzerrten und damit falschen Ergebnissen führen. Aus diesem Grund sind Ergebnisse der prospektiven Studien auf Stufe II und I glaubwürdiger als die der retrospektiven Studien.

In Kohortenstudien der Evidenzstufe II werden Therapie- und Kontrollgruppe von Beginn der Therapie an beobachtet und das

Eintreten von Zielereignissen in einem bestimmten Zeitraum erhoben. Die Therapie kann standardisiert werden und alle wichtigen Variablen, die den Krankheitsverlauf mit beeinflussen, können im Zeitverlauf direkt beobachtet werden. Bleibt die Einteilung eines bestimmten Patienten in die Therapie- bzw. die Kontrollgruppe dabei aber den Patienten oder ihren Ärzten überlassen, besteht die Gefahr, dass sich die beiden Gruppen von Anfang an hinsichtlich ihres Risikos für das Zielereignis unterscheiden. Z.B. wählen besonders vorsichtige Patienten eher die zusätzliche Therapie. Oder die Ärztin teilt Patienten, denen sie das exakte Einhalten des Therapieschemas nicht zutraut, bevorzugt der Kontrollgruppe zu. Am Ende bleibt die Frage offen, ob Unterschiede in den Ergebnissen auf die Therapie zurückgehen oder aber von vornherein durch die ungleiche Gruppeneinteilungen zu erwarten waren.

Dieses Problem des Selektions-Bias lösen die randomisierten Studien der Evidenzstufe I. In diesem Studiendesign wird die Zuordnung eines Studienteilnehmers zu einer der Gruppen allein durch den Zufall entschieden. Dadurch wird die Wahrscheinlichkeit maximiert, dass Patienten mit bekannten und unbekanntem Risikofaktoren für das Zielereignis gleichmäßig auf Therapie- und Kontrollgruppe verteilt werden und dadurch wirklich vergleichbare Gruppen entstehen.

Die CHARISMA-Studie, die Frau Dr. E. Vident zu Rate zieht, ist eine randomisiert kontrollierte Studie, ein RCT. Damit ist sie der Evidenzstufe I zugeordnet und besitzt aufgrund ihres Studientyps eine hohe Glaubwürdigkeit. Auch RCTs können Fehlerquellen beinhalten, die einen systematischen Fehler ermöglichen. Deswegen sollten weitere Details des Studiendesigns näher betrachtet werden, bevor einer Studie eine hohe Glaubwürdigkeit zugesprochen wird.

12.3.2 War die Studie randomisiert und verblindet?

Die Randomisierung sollte so durchgeführt werden, dass weder für den Arzt noch für den Patienten eine Wahlmöglichkeit bei der Gruppenzuteilung besteht. Dies geschieht z.B. bei einer Telefon-Randomisierung durch eine neutrale, räumlich getrennte Stelle. In der CHARISMA-Studie war eine solche externe Randomisierung gegeben. Unzulässige Wahlmöglichkeiten sind dagegen möglich, wenn die Patienten z.B. abwechselnd der Therapie- oder Kontrollgruppe zugeordnet werden, oder wenn anhand des Geburtsdatums oder des ersten Buchstaben des Namens entschieden wird. Häufig schließen die behandelnden Ärzte dann Patienten, die ihrer Meinung nach von der Therapie profitieren würden, nicht in die Studien ein, wenn sie offensichtlich der Kontrollgruppe zugeordnet werden würden. Ungewollt führt dieses Verhalten zu unterschiedlichen Risikokonstellationen in der Behandlungs- und der Kontrollgruppe.

Auch im weiteren Verlauf der Studie können Unterschiede zwischen Therapie- und Kontrollgruppe in der Wahrscheinlichkeit für das Zielereignis entstehen, wenn die Patienten oder ihre Behandler wissen, zu welcher Gruppe sie gehören. Bei einer möglicherweise risikoträchtigen Therapie werden die Behandler z.B. die Patienten der Therapiegruppe intensiver beobachten und negative Ereignisse zu einem früheren Zeitpunkt bemerken als bei den Patienten der Kontrollgruppe. Auch für die Patienten können Erwartungen oder Befürchtungen, die sie mit der Therapie verbinden, Auswirkungen auf das Eintreten des Zielereignisses haben. Unter dem Begriff Placeboeffekt versteht man z.B. die Beobachtung, dass Patienten, die eine Therapie wunschgemäß erhalten, grundsätzlich einen besseren Gesundheitszustand angeben als Patienten, die die Therapie nicht erhalten. Eine Lösung für diese Fehlerquel-

len ist die Verblindung der Studie. Man unterscheidet drei Typen der Verblindung:

- ▲ In einer einfach verblindeten Studie wissen die Patienten nicht, in welche Gruppe sie randomisiert wurden. In der Kontrollgruppe kommt häufig eine Placebotherapie zum Einsatz. Eine solche Placeboverblindung kann in einer Arzneimittelstudie besonders leicht durchgeführt werden.
- ▲ In einer doppelblinden Studie ist auch den Behandlern die Zuordnung ihrer Patienten nicht bekannt. Dadurch werden die oben beschriebenen Einflussmöglichkeiten des behandelnden Arztes ausgeschaltet.
- ▲ In einer dreifach verblindeten Studie wissen auch die Personen, die die Zielvariablen erheben nicht, zu welcher Gruppe der jeweilige Patient gehört.

In der CHARISMA-Studie war eine vollständige Verblindung der Patienten, der Behandler und der Erheber der Zielvariablen gegeben. Eine Ungleichbehandlung zwischen Patienten der Therapie- und der Kontrollgruppe zu irgendeinem Zeitpunkt der Studie ist damit unwahrscheinlich.

12.3.3 Waren die Gruppen zu Beginn der Studie vergleichbar?

Selbst in einer randomisierten Studie kann es vorkommen, dass die Studienteilnehmer in der Therapiegruppe sich zufällig in einzelnen Risikofaktoren für das Zielereignis von den Teilnehmern der Kontrollgruppe unterscheiden. Deswegen werden die bekannten Risikofaktoren in der Studie erhoben und im Ergebnisteil dargestellt. Für die CHARISMA-Studie sind diese Faktoren in der Tabelle 2 „Baseline Characteristics“ dargestellt. Für keines der dargestellten Merkmale gab es zu Beginn der Studie einen signifikanten Unterschied zwischen der Clopidogrel- und der

Placebogruppe. Die Randomisierung hat damit, zumindest für die bekannten Risikofaktoren der kardiovaskulären Zielereignisse, ihr Ziel erreicht. Sollten sich bei der Gegenüberstellung der Risikofaktoren zu Beginn der Studie Unterschiede zeigen, kann nicht ausgeschlossen werden, dass ein beobachteter Unterschied in den Zielereignissen nicht nur auf die untersuchte Therapie zurückgeht. Die Glaubwürdigkeit der Studienergebnisse wäre dann beeinträchtigt.

12.3.4 Schieden Patienten frühzeitig aus der Studie aus?

Prospektive Studien, die die positiven und negativen Folgen einer Therapie über mehrere Jahre beobachten, bieten die höchste Evidenzstufe für die Bewertung des Nutzens der Therapie. Ziel sollte dabei sein, das Schicksal aller Studienteilnehmer bis zum Ende der Beobachtungszeit zu verfolgen. Allerdings gibt es immer wieder Studienteilnehmer, die die Behandlung abbrechen, fortziehen oder aus anderen Gründen für die Erhebung der Zielvariablen nicht zur Verfügung stehen. Patienten, deren Zielereignisse nicht erhoben werden konnten, bezeichnet man als „lost to follow-up“. Sie können zu Verzerrungen des Studienergebnisses führen, wenn ihr Ausscheiden aus der Studie abhängig von ihrer Gruppenzuteilung ist. Es sind verschiedene Verzerrungsmöglichkeiten denkbar. Z.B. neigen gesunde Personen stärker dazu, ihren Wohnort zu wechseln. Führt in einer Studie die Therapie zu einer schnelleren Genesung als die Kontroll-Therapie, dann gehen in der Therapiegruppe mehr Personen durch Umzug verloren als in der Kontrollgruppe. Der positive Effekt der Therapie bildet sich in der Studie möglicherweise nicht adäquat ab.

Umgekehrt brechen in einer Studie, deren Therapie schwere Nebenwirkungen hervorruft, in der Therapiegruppe mehr Personen die Studienteilnahme ab als in der Kon-

trollgruppe. Ist der Abbruchgrund den Studienleitern nicht bekannt, unterschätzen sie möglicherweise die Nebenwirkungsrate in der Therapiegruppe, weil gerade die Personen mit schweren Nebenwirkungen für die Messung der Zielvariablen am Ende der Studie nicht mehr zur Verfügung stehen. Die Studienleiter sollten sich deswegen bemühen, das Schicksal ihrer Studienteilnehmer möglichst lückenlos zu verfolgen. Als Faustregel gilt: Bei einem Teilnehmerverlust von >20% in einer Studie können die Ergebnisse durch die Ausfälle erheblich verzerrt werden und müssen mit großer Skepsis betrachtet werden. Hochangesehene Journals verlangen sogar eine Nachbeobachtung von mehr als 90%. Sind in einer Studie Teilnehmerverluste vorgekommen, sollten diese genau beschrieben sein. Wie viele Verluste fielen in der Therapiegruppe, wie viele in der Kontrollgruppe an? Gibt es Hinweise darauf, dass sich die Ausfälle der beiden Behandlungsgruppen in ihrem Risiko für das Zielereignis unterscheiden? Die Auswirkungen der Ausfälle auf das Studienergebnis kann zusätzlich durch eine so genannte „Worst-case-Analyse“ eingeschätzt werden. Hierbei nimmt man an, dass alle verloren gegangenen Studienteilnehmern das negative Zielereignis erlitten haben und beobachtet, ob sich die Studienergebnisse unter dieser Konstellation deutlich verändern. Verändert sich dadurch das Ergebnis, indem sich z.B. ein vorher sichtbarer Unterschied zwischen den Gruppen relativiert oder sich sogar in das Gegenteil umkehrt, ist der Nutzen der Therapie nicht mit ausreichender Sicherheit belegt.

In der CHARISMA-Studie konnten nur 0,5% der Patienten der Therapiegruppe und 0,4% der Patienten der Kontrollgruppe nicht bis zum Studienende beobachtet werden, was eine ausgezeichnete Nachbeobachtung darstellt. Es ist nicht zu erwarten, dass die Ergebnisse der ausgefallenen Patienten das Studienergebnis deutlich verändert hätten.

12.3.5 Wechselten Patienten während der Studie zwischen den Studiengruppen?

Nicht immer führen Studienteilnehmer die Behandlung so durch, wie es die Studienleiter nach der Randomisierung in die Therapie- oder Kontrollgruppe vorgesehen haben. In der CHARISMA-Studie haben z.B. 20,4% der Patienten, die Aspirin und Clopidogrel einnahmen, und 18,2% der Patienten, die nur Aspirin einnahmen, die Therapie verändert oder abgebrochen. Es ist denkbar, dass ein Teil der Patienten der Therapiegruppe das Clopidogrel aufgrund der Nebenwirkungen nicht mehr einnahmen und damit dieselbe Therapie wie die Kontrollgruppe erhielten. Würde man diese Patienten bei der Analyse der Ergebnisse der Kontrollgruppe zuordnen („per protocol“- oder „as treated“-Analyse), wäre die ursprüngliche, mit der Randomisierung erzielte Gleichverteilung des Risikos für den Endpunkt nicht mehr gegeben. Es ist denkbar, dass Menschen, die die Therapie wegen ihrer Nebenwirkungen abbrechen, ein anderes Risiko für die Zielvariable haben, als Menschen, die die Therapie besser vertragen. Um ein verzerrtes Studienergebnis zu vermeiden, sollten alle Studienteilnehmer in der Gruppe analysiert werden, in die sie ursprünglich randomisiert wurden. Man spricht von einer „Intention-to-treat-Analyse“, die als Standardanalyse anerkannt ist. Die Therapie(un-)treue der Studienteilnehmer wird dann als Teil der Therapiefolgen betrachtet, sozusagen als Nebenwirkung der Therapie.

Die Intention-to-treat-Analyse ist immer ein Kompromiss, denn natürlich spiegelt auch sie nicht den wirklichen Effekt der Behandlung wider. Vielmehr stellt sie in der Regel eine konservative Einschätzung des Behandlungseffekts dar. Zum Schutze des Patienten hat man sich entschieden, lieber auf der konservativen Seite einen Irrtum zu begehen und den Behandlungseffekt geringer

einzuschätzen als er möglicherweise ist, als zu liberal eine Behandlung als wirksam zu bezeichnen, deren Effekt in Wirklichkeit gar nicht existiert. Ansonsten helfen nur eine maximale Nachbeobachtung und eine möglichst große Therapietreue, um den wirklichen Behandlungseffekt optimal zu schätzen.

In der CHARISMA-Studie wurde eine Intention-to-treat-Analyse durchgeführt. Die Zahl der Behandlungsabbrecher war insgesamt hoch und in der Therapiegruppe mit 20,4% signifikant höher als in der Kontrollgruppe mit 18,2% ($p < 0.001$). Es ist denkbar, dass sich ohne Intention-to-treat-Analyse

12.4 Was bedeuten die Ergebnisse der Studie?

Die CHARISMA-Studie hat sich im Verlauf der bisherigen Fragen als eine sinnvoll aufgebaute Studie erwiesen, deren Ergebnis glaubwürdig erscheint. Wie lautet aber das Ergebnis und wie kann es interpretiert werden? Hier soll die Bedeutung der wichtigsten statistischen Maßzahlen für den Behandlungseffekt am Beispiel der beiden primären Endpunkte für Wirksamkeit und unerwünschte Wirkungen der CHARISMA-Studie erklärt werden. Folgende Ergebnisse sind der Studie entnommen:

Tab. 12.2: Ergebnisse der CHARISMA-Studie

Zielvariable	Ereignisrate Clopidogrel + Aspirin (N=7 802)	Ereignisrate Placebo + Aspirin (N=7 801)	Relatives Risiko (RR)	95% CI
Primärer Endpunkt Wirksamkeit: Herzinfarkt oder Schlaganfall	534 (6,8%)	573 (7,3%)	0,93	[0,83–1,05]

ein höherer Nutzen der Therapie gezeigt hätte. Für Herrn Süßlich und Frau Dr. E. Vident ist nur das Ergebnis der Intention-to-treat-Analyse von Bedeutung. Zu dem Zeitpunkt, zu dem sie die Entscheidung für oder gegen die Einnahme von Clopidogrel zusätzlich zu Aspirin treffen müssen, wissen sie nicht, ob Herr Süßlich diese Therapie vertragen wird. Seine Situation entspricht der der Studienteilnehmer direkt nach der Randomisierung. Eine Einnahme von Clopidogrel zusätzlich zu Aspirin kann ihm nur empfohlen werden, wenn die Wirksamkeit für die Studienteilnehmer in dieser Situation nachgewiesen ist.

Insgesamt kommen Frau Dr. E. Vident und Herr Süßlich zu dem Schluss, dass die CHARISMA-Studie eine glaubwürdige Studie auf hohem Evidenzniveau ist. Es gibt keine Hinweise auf schwerwiegende Fehlerquellen, die das Ergebnis beeinträchtigen könnten. Mit diesem Wissen wenden sie sich den Ergebnissen der Studie zu.

Die statistischen Ergebniswerte beantworten grundsätzlich zwei Fragen:

- ▲ Wie groß ist der Behandlungseffekt?
- ▲ Wie groß ist der Unsicherheitsbereich (precision of the estimate) dieser Schätzung?

12.4.1 Schätzen des Behandlungseffektes

Die Ergebnisse einer Studie sind Häufigkeiten von Ereignissen wie Herzinfarkt oder Schlaganfall, die an den Personen einer Stichprobe beobachtet wurden. Diese gemessenen Ergebnisse werden als Schätzwerte für den wahren Wert in der Grundpopulation verwendet, an dem man eigentlich interessiert ist. Um diesen Schätzwert zu erhalten, werden im ersten Schritt die Ereignisraten in der Therapie- und Kontrollgruppe getrennt voneinander berechnet. Diese Ereignisraten beschreiben für die beiden Patientengruppen

jeweils das (absolute) Risiko, im Studienzeitraum ein Zielereignis zu erleiden.

Der primäre Endpunkt der CHARISMA-Studie ist ein zusammengesetzter Endpunkt aus den Ereignissen Herzinfarkt, Schlaganfall und kardiovaskulärer Tod. Wenn eines dieser drei Ereignisse im Beobachtungszeitraum von im Mittel 28 Monaten auftrat, galt der primäre Endpunkt als erreicht. Der primäre Endpunkt ist grundsätzlich der Endpunkt, für den die Studiengröße konzipiert wurde.

Die Ereignisrate der Clopidogrel/Aspirin-Gruppe (Experimental Event Rate=EER) beträgt für den kombinierten Endpunkt Herzinfarkt, Schlaganfall oder kardiovaskulärer Tod:

$$EER = \frac{\text{AnzahlEreignisse}}{\text{AnzahlPatienten}} = \frac{534}{7802} = 0,068 (6,8\%)$$

Die Ereignisrate der Placebo-Gruppe (Control Event Rate=CER) beträgt entsprechend:

$$CER = \frac{\text{AnzahlEreignisse}}{\text{AnzahlPatienten}} = \frac{573}{7801} = 0,073 (7,3\%)$$

Die prognostizierte Wirksamkeit der untersuchten Therapie sollte zu einem Unterschied dieser Ergebnisraten zu Gunsten der Clopidogrel/Aspirin-Therapie führen. Um diesen Unterschied darzustellen, gibt es eine Reihe von sinnvollen Maßzahlen. Dabei unterscheidet man zwischen den Relativen Risikoschätzern, bei denen man die Ereignisrate der Behandlungsgruppe zu der Ereignisrate der Kontrollgruppe in Beziehung stellt ($EER_{\text{Behandlungsgruppe}}$ wird durch die $EER_{\text{Kontrollgruppe}}$ geteilt) und die Absoluten Risikoschätzer, bei der die Ereignisrate_{Behandlungsgruppe} von der Ereignisrate_{Kontrollgruppe} abgezogen wird. Die wichtigsten Risikoschätzer sind das Relative Risiko, bzw. die Relative Risiko-Reduktion und die Absolute Risiko-Reduktion und ihr Kehrwert, die „Number needed to treat“.

Die Bedeutung dieser unterschiedlichen Risikoschätzer lassen sich grafisch folgendermaßen darstellen:

Tab. 12.3: Die wichtigsten Risikoschätzer

Relatives Risiko:	Relative Risiko-Reduktion:	Absolute Risiko-Reduktion:	Number needed to treat:
$RR = \frac{EER}{CER}$	$RRR = 1 - RR$	$ARR = CER - EER$	$NNT = \frac{1}{ARR}$
$RR = \frac{0,068}{0,073} = 0,9$	$RRR = 1 - 0,93 = 0,0$	$ARR = 0,073 - 0,068 = 0,0$	$NNT = \frac{1}{0,005} = 20$

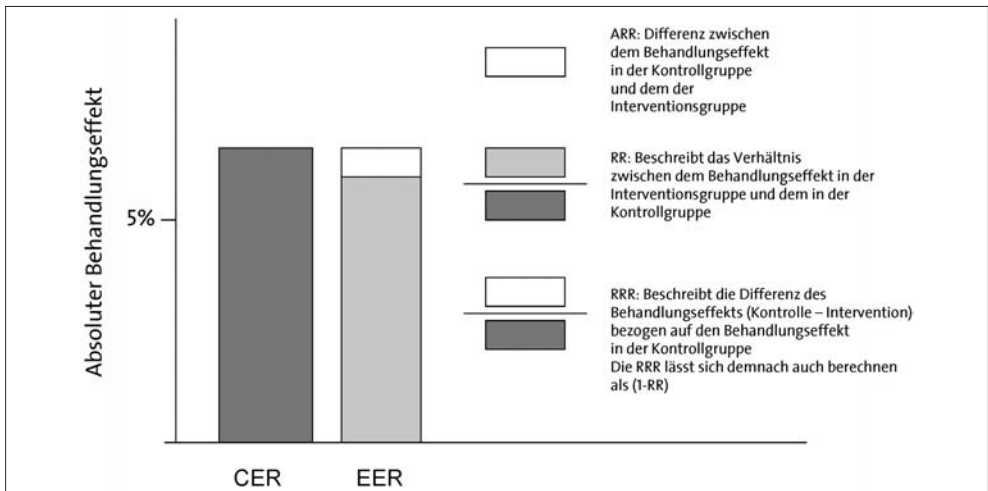


Abb. 12.2: Relativer und Absoluter Risikoschätzer bei binären Variablen

In der obigen schematischen Abbildung stellt die dunkelgraue Säule das Absolute Risiko (Inzidenz) der Placebogruppe und die hellgraue Säule das Absolute Risiko der Therapiegruppe dar. Das Relative Risiko ist das Verhältnis dieser beiden Werte zueinander und beantwortet die Frage, um **welchen Faktor** das Risiko in der Therapiegruppe niedriger ist als in der Placebogruppe. Die Relative Risiko-Reduktion ist eine Maßzahl für den Anteil, um den sich das Absolute Risiko der Placebogruppe durch die Therapie senken lässt. Die Absolute Risiko-Reduktion beschreibt dagegen den absoluten Unterschied zwischen dem Absoluten Risiko in der Therapie- und Placebogruppe.

12.4.2 Was bedeuten die Ergebnisse der CHARISMA-Studie?

RR (Relatives Risiko) = 0,93

Die Ereignisrate in der Clopidogrel-Gruppe ist um den Faktor 0,93 geringer als die Ereignisrate in der Placebogruppe. Relative Risiken in Therapiestudien können, wenn das Zielereignis ein unerwünschtes Ereignis ist, folgendermaßen interpretiert werden:

- ▲ Hat die Therapie keinen Effekt auf die Ereignisraten, dann sind die Ereignisraten der beiden Gruppen gleich groß. Das RR liegt bei 1.
- ▲ Hat die Therapie einen positiven Effekt, liegt die Ereignisrate in der Therapiegruppe unter der Ereignisrate der Kontrollgruppe. In diesem Fall liegt das Relative Risiko zwischen 0 und 1.
- ▲ Hat die Therapie dagegen einen negativen Effekt, liegt die Ereignisrate der Therapiegruppe höher als die Ereignisrate der Kontrollgruppe. Das Relative Risiko ist in diesem Fall größer als 1.

Das Relative Risiko in der Studie liegt unter 1, also hat die Clopidogrel/Aspirin-Therapie grundsätzlich einen positiven Effekt auf die Ereignisrate.

RRR (Relative Risiko-Reduktion) = 0,07

Durch die Einnahme der Kombination Clopidogrel/Aspirin konnte die Ereignisrate im Vergleich zur Einnahme von Aspirin allein um 7% gesenkt werden.

Ob eine Senkung der Ereignisrate um 7% klinisch relevant ist, kann aus einem Relativen Risikoschätzer allein nicht geschlossen werden. In der Regel erwartet man von neuen Behandlungen für einen klinisch relevanten Effekt eine Wirksamkeit im Bereich von 20%–30% weniger unerwünschter Ereignisse wie Herzinfarkt oder Schlaganfall. Möglicherweise kann für ein häufiges Zielereignis aber auch eine Absenkung um 7% bereits eine bedeutsame Anzahl von Ereignissen verhindern. Aus diesem Grund sollte zur Beurteilung der klinischen Relevanz eines Ergebnisses immer die Absolute Risiko-Reduktion oder die „Number needed to treat“ herangezogen werden.

ARR (Absolute Risiko-Reduktion) = 0,005

Durch die Therapie mit Clopidogrel zusätzlich zu Aspirin kann das Auftreten eines Zielereignisses (Herzinfarkt oder Schlaganfall) um 0,5% (oder genauer gesagt: um 0.5 Prozentpunkte) gesenkt werden. Mit anderen Worten, das Ereignis kann bei 5 von 1000 behandelten Personen verhindert werden. Im Gegensatz zum Relativen Risiko oder der Relativen Risiko-Reduktion kommt bei der Absoluten Risiko-Reduktion das Ausgangsrisiko der Kontrollgruppe, ein unerwünschtes Ereignis zu erleiden, zur Geltung. Aus diesem Grund reflektiert die Absolute Risiko-Reduktion den wirklichen Nutzen, den ein Patient von einer Behandlung erwarten kann.

Die bisher verwendeten Risiko-Maßeinheiten sind relativ abstrakt und Ärzten und Patienten in der Praxis oft nur schwer zu vermitteln. Eine intuitiv eingängigere Bedeutung kommt dem mathematischen Kehrwert der Absoluten Risiko-Reduktion zu, der so genannten Number needed to treat.

NNT (number needed to treat to prevent one additional event) = 200

Es müssen 200 Patienten mit Clopidogrel/Aspirin behandelt werden, um innerhalb von fünf Jahren ein zusätzliches Zielereignis zu verhindern. Die 199 anderen Patienten haben keinen Nutzen von ihrer Behandlung.

Aus den hier vorgestellten Punktschätzern für die Wirksamkeit der Kombination Clopidogrel/Aspirin zur Vermeidung von Herzinfarkten und Schlaganfällen kann eine gewisse Wirksamkeit für die neue Therapie abgeleitet werden. Zu Beginn der Studie hatten die Forscher jedoch festgelegt, dass der kleinste für sie noch klinisch relevante Effekt bei einer Risiko-Reduktion von 20% liegen würde. Der beobachtete Effekt ist deutlich kleiner als die postulierte 20% Relative Risiko-Reduktion. Die Tatsache, dass von 200 Patienten, die das Medikament fünf Jahre lang einnehmen, nur einer einen Nutzen davon hat, unterstreicht die geringe Bedeutung der Wirksamkeit.

Die unterschiedliche Bedeutung von Relativen und Absoluten Schätzern für den Unterschied zwischen Therapie und Placebogruppe wird an folgendem konstruierten Beispiel deutlich:

In einer Stichprobe von 1000 Personen mit akuter Tonsillitis wurden 500 Personen

mit einem Antibiotikum behandelt, 500 Personen erhielten ein Placebo-Medikament. Nach acht Tagen Behandlungszeit wird der negative Endpunkt „Schmerzen noch vorhanden“ überprüft.

1. Szenarium In der Antibiotikum-Gruppe haben 50 Personen acht Tage nach Behandlungsbeginn noch Schmerzen, in der Placebo-Gruppe sind es 100 Personen.

2. Szenarium: In der Antibiotikum-Gruppe haben 200 Personen acht Tage nach Behandlungsbeginn noch Schmerzen, in der Placebo-Gruppe sind es 400 Personen.

In beiden dargestellten Szenarien ist das Relative Risiko 0,5, also die Ereignisrate in der Therapiegruppe nur halb so groß wie in der Placebogruppe. Weil aber das Ereignis, nach acht Tagen noch Schmerzen zu haben, im 1. Szenarium deutlich seltener auftritt als im zweiten Szenarium, ist die Number needed to treat entsprechend höher: Im ersten Szenarium müssten 10 Personen das Antibiotikum erhalten, damit eine Person nach acht Tagen zusätzlich schmerzfrei ist, im zweiten Szenarium hätte bereits eine von 3 ($\approx 2,5$) Personen diesen Vorteil von der Behandlung.

Tab. 12.4: Bedeutung von Relativen und Absoluten Schätzern für den Unterschied zwischen Therapie- und Placebogruppe

Maßzahl	1. Szenarium		2. Szenarium	
	Antibiotikum-Gruppe N=500	Placebogruppe N=500	Antibiotikum-Gruppe N=500	Placebogruppe N=500
Absolutes Risiko, nach acht Tagen noch Schmerzen zu haben	$EER = \frac{50}{500} = 0,1$	$CER = \frac{100}{500} = 0,2$	$EER = \frac{200}{500} = 0,4$	$CER = \frac{400}{500} = 0,8$
Relatives Risiko	$RR = \frac{0,1}{0,2} = 0,5$		$RR = \frac{0,4}{0,8} = 0,5$	
Relative Risiko-Reduktion	$ARR = 0,2 - 0,1 = 0,1$		$ARR = 0,8 - 0,4 = 0,4$	
Number needed to treat	$NNT = \frac{1}{0,1} = 10$		$NNT = \frac{1}{0,4} = 2,5$	

12.4.3 Schätzen der Unsicherheit

Die Ergebnisse einer klinischen Studie werden als Schätzwerte für den „wahren“ Effekt der Therapie in der Grundpopulation angesehen. Die in der CHARISMA-Studie gemessenen Werte zeigen einen, wenn auch geringen, Unterschied in der Ereignisrate der Therapie- und Kontrollgruppe. In der Clopidogrel-Gruppe kamen 7% weniger Schlaganfälle und Herzinfarkte vor als in der Placebo-Gruppe. Dieser Unterschied kann entweder einen tatsächlich vorhandenen Unterschied in der Grundgesamtheit abbilden, oder aber zufällig durch eine (unglückliche) Auswahl der Patienten entstanden sein. Konfidenzintervalle und statistische Tests sind Werkzeuge zur Einschätzung der Wahrscheinlichkeit eines solchen Irrtums.

12.4.4 Konfidenz- oder Vertrauensintervall

Heute werden Relative Risikoschätzer fast immer mit einem Konfidenzintervall angegeben. Das Konfidenzintervall ist eine Information darüber, in welchem Zahlenbereich der wahre Wert der Grundpopulation mit hoher Wahrscheinlichkeit zu erwarten ist. Das 95%-Konfidenzintervall für das Relative Risiko des primären Endpunktes in der CHARISMA-Studie schließt den Zahlenbereich von 0,83 bis 1,05 ein. Hieraus lassen sich zwei Informationen ableiten:

- ▲ Mit 95% Wahrscheinlichkeit liegt der „wahre Wert“ des Relativen Risikos in der Grundpopulation zwischen 0,83 und 1,05, d.h. mit diesem Konfidenzintervall ist vereinbar, dass im besten Falle die Kombinationstherapie in der Grundpopulation als wahren Wert eine Relative Risiko-Reduktion von 17% für Herzinfarkt/Schlaganfall erzielt, im ungünstigsten Falle aber sogar ein Anstieg der Herzinfarkt-/Schlaganfallrate um 5% verursacht.

- ▲ Das Konfidenzintervall schließt die 1 als neutrales Relatives Risiko ein, der beobachtete Unterschied von 7% ist damit nicht signifikant auf dem 5% Niveau.

Das Konfidenzintervall schließt mit dem zweiten Punkt die Information des p-Wertes mit ein, der im nächsten Abschnitt betrachtet wird. Liegt das 95%-Konfidenzintervall eines Relativen Risikos vollständig über oder unter 1, ist der beobachtete Unterschied zwischen den beiden Gruppen signifikant. Grundsätzlich kann ein Relativer oder Absoluter Risikoschätzer nur in Verbindung mit seinem Konfidenzintervall sinnvoll interpretiert werden.

Um die Bedeutung des Konfidenzintervalls deutlich zu machen, werden hier neben dem Endpunkt kardiovaskuläres Ereignis auch die Ergebnisse der Studie für die beiden Endpunkte „schwere Blutung“ und „moderate Blutung“ betrachtet.

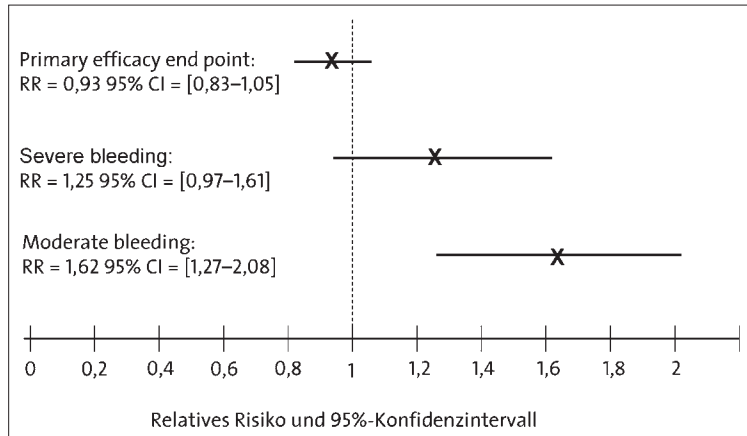
Abbildung 12.3 stellt die drei Relativen Risikoschätzer mit ihren 95%-Konfidenzintervallen grafisch dar.

Aus der Abbildung wird ersichtlich, dass die Konfidenzintervalle der Relativen Risikoschätzer für die Endpunkte „Primary efficacy endpoint“ und „Severe bleeding“ die 1 als neutralen Wert überlappen. Diese beiden Unterschiede sind nicht signifikant, d.h. die Wahrscheinlichkeit, dass die Unterschiede in der Grundpopulation nicht vorhanden sind, ist größer als 5%. Das Konfidenzintervall für den Relativen Risikoschätzer des Endpunktes „Moderate bleeding“ liegt dagegen vollständig über 1. Dieses Ergebnis ist signifikant. Für alle drei Schätzer gilt: Der wahre Wert der Grundpopulation liegt mit 95% Wahrscheinlichkeit innerhalb des dargestellten Konfidenzintervalls.

12.4.5 Statistischer Test

In den meisten Studien wird neben dem Konfidenzintervall auch der so genannte p-

Abb. 12.3: Relatives Risiko und 95%-Konfidenzintervall für drei Endpunkte der CHARISMA-Studie



Wert angegeben, obwohl die Information, die er beinhaltet, bereits aus dem Konfidenzintervall entnommen werden kann. Mit Hilfe eines statistischen Tests wird die Wahrscheinlichkeit untersucht, dass ein in einer Stichprobe beobachteter Unterschied in Wirklichkeit in der Grundgesamtheit nicht vorhanden ist (Fehler 1. Art). Im Vorfeld der Studie wird die akzeptierte Irrtumswahrscheinlichkeit festgelegt. In der Regel ist es das 5% Niveau für einen Fehler 1. Art. Ein signifikanter Unterschied in der Ereignisrate der beiden Gruppen bedeutet dann:

Mit 95% Wahrscheinlichkeit ist auch in der Grundgesamtheit ein Unterschied zwischen den Gruppen vorhanden. Mit anderen Worten, wenn man 100 Stichproben mit dem beobachteten Ergebnis hätte, könnte man davon ausgehen, dass hinter 95 dieser Stichproben eine Grundgesamtheit mit einem Unterschied liegt, 5 der Stichproben kämen aus Grundgesamtheiten, in denen kein Unterschied vorliegt.

Je nach Messniveau der Endpunkte und deren Verteilungen kommen unterschiedliche statistische Tests zur Anwendung. Beim Vergleich zweier Gruppen hinsichtlich der Häufigkeit eines binären Zielereignisses wird in der Regel der χ^2 -Test angewandt.

In der CHARISMA-Studie wurde für den Vergleich der primären Endpunkte ein χ^2 -Test durchgeführt. Der Unterschied wäre sig-

nifikant auf dem 5%-Niveau, wenn der p-Wert unter 0,05 läge. Der p-Wert in der Studie ist aber 0,22 und liegt damit deutlich über dieser Grenze. Der Unterschied zwischen den Ereignisraten der Clopidogrel-Gruppe und der Placebogruppe ist damit nicht signifikant. Es kann also nicht mit 95% Wahrscheinlichkeit angenommen werden, dass sich auch in der Grundpopulation ein Unterschied zwischen „mit Clopidogrel/Aspirin“ versus „nur Aspirin“ existiert.

12.5 Zusammenfassung der evidenzbasierten Studienbewertung

Für Frau Dr. E. Vident und Herrn Süßlich bedeuten die Ergebnisse der CHARISMA-Studie:

In der Studie wurde ein geringer Unterschied in den Herzinfarkt- und Schlaganfallraten zugunsten der Kombinationsgruppe Clopidogrel/Aspirin beobachtet. Das Relative Risiko für diesen Endpunkt war 0,93, das 95%-Konfidenzintervall betrug [0,83–1,05]. Das Ergebnis lässt die Möglichkeit offen, dass die zusätzliche Therapie mit Clopidogrel keinen Einfluss auf das Risiko für Herzinfarkt und Schlaganfall hat oder das Risiko möglicherweise sogar erhöht. Die Überlegenheit der Kombinationstherapie Clopidogrel/Aspirin ist damit nicht belegt. In diesem

Fall ist es nicht nötig, zusätzlich eine Number needed to treat zu berechnen. In der CHARISMA-Studie wurde diese Maßzahl auch nicht angegeben.

Auch in der Häufigkeit der schweren Blutungen als wichtigste unerwünschte Nebenwirkungen wurde in der Studie ein Unterschied zwischen Therapie- und Kontrollgruppe beobachtet. Das Relative Risiko für diesen Endpunkt war 1,25, das 95%-Konfidenzintervall betrug [0,97–1,61], d.h. in der Clopidogrel/Aspirin-Gruppe traten schwere Blutungen um 25% häufiger auf als in der Kontrollgruppe. Auch für dieses Ergebnis schließt das Konfidenzintervall die 1 ein und es kann damit nicht ausgeschlossen werden, dass auch für die unerwünschte Nebenwirkung in der Grundpopulation kein Unterschied zwischen den beiden Gruppen besteht. Allerdings ist man beim Auftreten von Nebenwirkungen sensibler. Zugunsten der Sicherheit des Patienten nimmt man bereits einen nicht signifikanten Trend für vermehrte Nebenwirkungen in der Studienpopulation ernst. Das bedeutet, man ist eher bereit, bei den Schlussfolgerungen auf die Grundpopulation den Irrtum einzugehen, dass Nebenwirkungen auftreten, wenn diese möglicherweise gar nicht vorkommen, als das Risiko einzugehen, fälschlicherweise zu folgern, dass Nebenwirkungen nicht existieren, wenn sie in Wirklichkeit in der Grundpopulation auftreten.

Die CHARISMA-Studie wurde auf Evidenzstufe I angesiedelt und wies keine schwerwiegenden Mängel auf, die einen systematischen Fehler wahrscheinlich machen. Die Ergebnisse sind grundsätzlich glaubwürdig. Die gemessene Wirksamkeit der Kombinationstherapie Clopidogrel/Aspirin ist gering und ein Zufallsbefund (Zufallsfehler) kann aufgrund der beobachteten Werte nicht ausgeschlossen werden. Die Wirksamkeit der Kombination Clopidogrel/Aspirin im Vergleich zur Monotherapie Aspirin zur Verhinderung von Herzinfarkt und Schlaganfall ist in dieser Studie nicht ausreichend nachgewiesen.

12.5.1 Sind die Ergebnisse auf den zu behandelnden Patienten übertragbar?

Nachdem die Studie bewertet und ihre Glaubwürdigkeit eingeschätzt wurde, müssen Frau Dr. E. Vident und Herr Süßlich entscheiden, ob die Ergebnisse auf Herrn Süßlichs Situation übertragbar sind. Patienten in klinischen Studien unterscheiden sich häufig in unterschiedlichen Parametern von dem Durchschnitt der Patienten mit der entsprechenden Indikation. Ältere, multimorbide Menschen werden oft nicht in die Studien eingeschlossen. Auch Schwangere und Kinder finden sich kaum in den Studienpopulationen. Auch wenn die Gründe für den Ausschluss besonders gefährdeter Gruppen einleuchten – die Übertragung der Studienergebnisse auf diese Gruppen ist entsprechend erschwert.

Grundsätzlich wird in einer Studie von einer Stichprobe auf eine Grundpopulation geschlossen. Die Ergebnisschätzer gelten damit nur für Grundpopulationen, für die die Stichprobe repräsentativ ist. Jedes Mitglied der Grundpopulation muss mit der gleichen Wahrscheinlichkeit Teilnehmer der Studie werden können. Die Übertragung der geschätzten Werte auf einen einzelnen Patienten ist im zweiten Schritt umso besser möglich, wenn die gesundheitliche Situation des Patienten der Grundpopulation möglichst gut entspricht.

Herr Süßlich trägt mehrere Risikofaktoren für eine koronare Herzkrankheit und wäre daher als Teilnehmer der Studie in Frage gekommen. Er leidet unter einem insulinpflichtigen Diabetes, genau wie 81% der Studienteilnehmer. Dazu kommt ein seit mehr als drei Monaten bestehender Bluthochdruck, genau wie bei 47% der Studienteilnehmer. Zusätzlich gibt es Hinweise auf eine koronare Herzerkrankung, wie sie bei 78% der Teilnehmer diagnostiziert wurde. Er ist mit 57 Jahren etwas jünger als der Durch-

schnitt der Studienteilnehmer (64 Jahre), erfüllt aber das Einschlusskriterium (>45 Jahre). Ebenso wie Herr Süßlich sind 70% der Studienteilnehmer männlichen Geschlechts. Soweit bekannt, liegen bei Herrn Süßlich keine unmittelbaren Gründe für eine Clopidogrel-Therapie vor und auch die anderen Ausschlussgründe der Studie treffen nicht auf ihn zu. Aufgrund der Ein- und Ausschlusskriterien kann man also davon ausgehen, dass die Studienergebnisse auf Herrn Süßlich übertragbar sind. Die in Tabelle 12.2 dargestellten Ergebnisse der Untergruppenanalyse zeigen zusätzlich, dass auch für Studienteilnehmer mit Herrn Süßlichs spezieller Situation (z.B. für Diabetiker oder für Männer allein) keine signifikante Wirksamkeit der Clopidogrel/Aspirin-Therapie nachgewiesen wurde.

12.5.2 Waren die Endpunkte patientenrelevant?

Primärer Endpunkt für die Wirksamkeit der CHARISMA-Studie war der kombinierte Endpunkt Herzinfarkt, Schlaganfall oder Tod aufgrund kardiovaskulärer Ursachen. Eines dieser Ereignisse wäre für Herrn Süßlich hochgradig relevant, weil es sein Leben und seine Lebensqualität stark beeinflussen würde. Zwar wird aus der Beschreibung der Endpunkte im Methodenteil der Studie nicht ausreichend deutlich, ob auch gesundheitlich folgenlose Schlaganfälle eingeschlossen waren. Allerdings zeigt die Aufstellung der unterschiedlichen Endpunkte in Tabelle 4, dass auch für die Ereignisse mit Todesfolgen kein Unterschied zwischen den beiden Behandlungsgruppen bestand.

Problematisch sind Studien, die so genannte Surrogat-Endpunkte verwenden. Wenn die Wirksamkeit einer Therapie ausschließlich durch Laborparameter nachgewiesen wird, ohne dass diese sich zwangsläufig auf Gesundheit und Lebensqualität der

Patienten auswirken, sollten weitere Studien mit patientenrelevanten Endpunkten herangezogen werden.

12.5.3 Wie stehen positive und negative Folgen der Therapie zueinander in Beziehung?

Wie die meisten Therapien und präventiven Maßnahmen hat die zusätzliche Behandlung mit Clopidogrel zu Aspirin nicht nur positive Wirkungen, sondern auch unerwünschte Nebenwirkungen. Wäre der Unterschied hinsichtlich der primären Zielereignisse signifikant, müsste diese Wirksamkeit gegen den Schaden der Therapie abgewogen werden. Im Falle der CHARISMA-Studie stehen einer Risikoabsenkung für Herzinfarkt und Schlaganfall von 7% ein Anstieg des Risikos für schwere Blutungen um 25% entgegen. Allerdings treten die Blutungen seltener auf als die Zielereignisse Herzinfarkt und Schlaganfall. Aus der Absoluten Risiko-Reduktion kann abgeleitet werden, dass bei 5 von 1000 über fünf Jahre mit Clopidogrel zusätzlich zu Aspirin behandelten Personen ein kardiovaskuläres Ereignis verhindert werden kann. Gleichzeitig käme es bei 4 Personen zusätzlich zu schweren Blutungen. Wirksamkeit und Schaden der Therapie liegen damit eng beieinander, sodass allein aufgrund dieser Zahlen nicht von einem eindeutigen Nutzen der Therapie gesprochen werden kann. Dazu kommt natürlich, dass die Ergebnisse nicht signifikant sind und auch zufällig durch die Auswahl der Stichprobe entstanden sein können.

Frau Dr. E. Vident und Herr Süßlich diskutieren das Ergebnis der Studie im Rahmen ihres nächsten Arzt-Patienten-Gesprächs und kommen zu der Entscheidung, dass die zusätzliche Clopidogrel-Therapie für Herrn Süßlich in seiner momentanen Situation nicht sinnvoll ist. Allerdings fragt Herr Süßlich seine Ärztin, ob die Wirksamkeit von vermehrter Bewegung, die sie

ihm so dringend ans Herz lege, denn in Studien bewiesen sei. Frau Dr. Evident nimmt diese Frage in den nächsten Qualitätszirkel der Hausärzte mit ...

Literatur

- [1] Bhatt DL et al., Clopidogrel and Aspirin versus Aspirin Alone for the Prevention of Atherothrombotic Events. N Engl J Med. (2006), 1706–1717

13 Kritische Bewertung von systematischen Reviews und Meta-Analysen

Heiner C. Bucher

Systematische Übersichtsarbeiten und Meta-Analysen von Studien zum Nutzen von Behandlungsmaßnahmen sind wichtige Informationsquellen für Kliniker. Viele Behandlungsrichtlinien stützen sich auf Ergebnisse von Meta-Analysen. Die **Cochrane Library** fasst in systematischen Übersichtsarbeiten und Meta-Analysen die Evidenz zu therapeutischen Maßnahmen aus randomisierten kontrollierten Studien aus allen Bereichen der Medizin zusammen.

Im Gegensatz zu unsystematischen Übersichtsarbeiten oder durch Experten verfassten Zusammenfassungen weisen systematische Übersichtsarbeiten mehrere Vorteile auf. Bei klar formulierten Einschlusskriterien der berücksichtigten Studien bieten Meta-Analysen reproduzierbare Informationen über den Nutzen von therapeutischen Interventionen. Informationsflut und Zeitmangel zwingen Ärztinnen und Ärzte zu einem rationalen Zeitmanagement in der Informationsbeschaffung und Fortbildung. Übersichtsarbeiten sind deshalb eine willkommene und populäre Informationsquelle von klinisch tätigen Ärztinnen und Ärzten. Es zeigt sich jedoch, dass unsystematische Übersichtsarbeiten und medizinische Textbücher, in welchen Meinungen von Autoren und Studieninterpretationen vermischt werden, häufig nicht auf dem neuesten wissenschaftlichen Stand sind oder einseitige Gewichtungen der Evidenz von (positiven) Studienergebnissen vornehmen [1; 2].

Die Meta-Analyse ist eine Methode, Studien mittels systematischer computerisierter Literatursuche zu identifizieren und die Ergebnisse mit statistischen Methoden zu kombinieren und zu quantifizieren (s. Tab.

13.1). Eine Meta-Analyse kann insbesondere dann wertvolle Hinweise über einen Behandlungsnutzen geben, wenn die Ergebnisse von mehreren kleineren Studien keine zuverlässigen Schlüsse zulassen. Dies kann dann eintreten, wenn Ergebnisse einzelner Studien widersprüchlich sind oder wenn aufgrund kleiner Stichproben statistisch nicht signifikante Resultate vorliegen.

Lau und Mitarbeiter zeigten beispielsweise, dass der Nutzen der Thrombolyse beim akuten Myokardinfarkt lange Zeit verkannt wurde [3]. Bereits zu Beginn der 80er-Jahre zeigten Meta-Analysen, dass diese Behandlung die Herzinfarkt mortalität statistisch signifikant senkt. Die großen, Ende der 80er-Jahre publizierten Studien bestätigten diese Ergebnisse und das in der Meta-Analyse gefundene Ausmaß der Mortalitätsreduktion. Mit anderen Worten: Der Einbezug von Meta-Analysen in die klinische Forschung und Praxis kann zu einem früheren Zeitpunkt Informationen über das Ausmaß des Nutzens oder Schadens einer Intervention liefern.

Meta-Analysen können jedoch auch zu falschen Schlussfolgerungen führen. Die empirische Forschung zeigt, dass in mehreren Fällen Ergebnisse von Meta-Analysen, die sich auf Studien mit kleinen Fallzahlen stützten, in nachfolgenden klinischen Studien mit großen Fallzahlen nicht bestätigt werden konnten [4]. Zum Beispiel fand eine Meta-Analyse, dass Magnesium zu einer statistisch signifikanten Reduktion der Mortalität bei akutem Myokardinfarkt führt [5]. Die nachfolgende ISIS-4 Studie [6], die über 55 000 Patienten einschloss, konnte diesen Effekt

Tab 13.1: Kriterien zur Beurteilung der Methodik einer systematischen Übersichtsarbeit (Meta-Analyse) [8]

1. Sind die Ergebnisse der Studie glaubwürdig?
<ul style="list-style-type: none"> • Behandelt die Übersichtsarbeit eine genau umschriebene klinische Fragestellung? • Sind die Ein- und Ausschlusskriterien, nach denen Studien berücksichtigt wurden, detailliert aufgeführt? • Wie wahrscheinlich ist es, dass relevante Studien nicht berücksichtigt wurden (<i>Publication bias</i>)? • Welche Suchstrategien wurden verwendet, um unveröffentlichte Studien zu identifizieren? • Wurde die Validität der eingeschlossenen Studien überprüft? • Ist die Art, wie die Studien überprüft wurden, reproduzierbar?
2. Wie lauten die Studienergebnisse?
<ul style="list-style-type: none"> • Was sind die Gesamtergebnisse der Meta-Analyse? • Wie exakt sind die Ergebnisse (Vertrauensintervalle)? • Sind die Studienergebnisse in den einzelnen Studien vergleichbar?
3. Sind die Ergebnisse für die Behandlung meiner Patienten nützlich?
<ul style="list-style-type: none"> • Können die Ergebnisse auf den/die in meiner Behandlung stehenden Patienten/in übertragen werden? • Wurden alle relevanten klinischen Endpunkte in der Analyse berücksichtigt? • Überwiegen die positiven Auswirkungen die schädlichen Auswirkungen und Kosten der Maßnahme?

nicht bestätigen. Die Meta-Analyse schloss eine überwiegend große Zahl von kleinen Studien mit großem Behandlungsnutzen ein, Studien mit negativen Ergebnissen fehlten. Der wahrscheinliche Grund des diskrepanten Ergebnisses ist der so genannte Publication-Bias [7]. Studien mit negativen Ergebnissen werden weniger häufig publiziert und können deshalb, falls sie nicht identifiziert und in die Analyse einbezogen werden, das Ergebnis einer Meta-Analyse verfälschen.

Die Ergebnisse einer Meta-Analyse sollten deshalb kritisch interpretiert werden, wie nachfolgend anhand eines konkreten Fallbeispiels gezeigt wird [8].

13.1 Klinisches Szenario

Eine 35-jährige Asylbewerberin aus Zaire konsultiert Ihre Praxis. Sie beklagt seit ihrer Einreise vor fünf Monaten drei Kilo Gewichtsverlust und rezidivierende Infekte der oberen Luftwege. In der klinischen Untersuchung fällt Ihnen eine Candida-Stomatitis auf und Sie haben den Verdacht auf eine HIV-Infektion. Die Patientin willigt zu einem HIV-Test ein, der positiv ausfällt. Die weite-

ren Analysen zeigen eine mittelgradige Immunsuppression mit einer CD4-Zellzahl von 350 per mm³ und eine HIV-1-RNA-Viruslast von 35 000 Kopien pro ml. Ein intrakutaner Tuberculintest zeigt eine Erhabenheit von 10 mm. Eine antiretrovirale Therapie wäre indiziert, doch ob und wann sie eine Kostengutsprache von den Asylbehörden erhalten wird, ist noch unklar. Sie überlegen sich, ob die Patientin von einer medikamentösen Tuberkulose-Prophylaxe profitiert, sind jedoch unsicher. Aufgrund des normalen Thoraxröntgenbildes ist eine aktive Tuberkulose sehr unwahrscheinlich. Sie beschließen, weitere Information zu beschaffen. Ihre Fragestellung ist: Was ist der Nutzen einer Tuberkulose-Prophylaxe bei HIV-Infektion und wie sind die zu erwartenden Nebenwirkungen?

13.2 Die Evidenz suchen

Sie suchen mit PubMed im Internet und geben als Suchbegriffe (medical subject headings, so genannte MESH) den Begriff „tuberculosis“ ein. Das Indexsystems verweist Sie auf den Unterbegriff „antitubercular agents“. Sie machen eine Suche mit diesem Unterbe-

griff und schränken diese mit dem Suchbefehl „Publikationstyp meta-analysis“ ein. Sie identifizieren zwei Meta-Analysen und entscheiden sich für die Arbeit neueren Datums über Isoniazid-Prophylaxe, der primär empfohlenen Standardtherapie für die prophylaktischen Behandlung der latenten Tuberkulose bei HIV-Infizierten. Die Arbeit ist im Archiv des Journals im Volltext verfügbar (<http://www.aidsonline.com>).

13.3 Die Evidenz bewerten

Kritische Evaluation einer systematischen Übersichtsarbeit

Meta-Analysen sind methodisch anspruchsvoll und können zu fehlerhaften Schlussfolgerungen führen. Die Fähigkeit zur kritischen Beurteilung der Qualität und klinischen Brauchbarkeit einer Meta-Analyse sind deshalb wichtig für eine korrekte klinische Entscheidungsfindung. Die Validität einer Meta-Analyse und die Applizierbarkeit der Ergebnisse auf das konkrete klinische Problem kann anhand der folgenden Kriterien überprüft werden [8] (s. Tab. 13.1). Wir beziehen uns im Folgenden auf die Meta-Analyse unseres Beispiels [9].

13.3.1 Sind die Ergebnisse der Studie glaubwürdig?

Behandelt die Übersichtsarbeit eine umschriebene klinische Fragestellung?

Viele nicht systematische Übersichtsarbeiten vermitteln Informationen zu verschiedenen Bereichen eines klinischen Problems wie Epidemiologie, Pathogenese, Diagnose, Behandlung und Prognose. Derartige Informationen sind als Einführung nützlich, vielfach bleiben jedoch Schlussfolgerungen zu praktischen Belangen vage. Eine Meta-Analyse sollte eine konkrete klinische Fragestellung beinhalten (vgl. Kap. 8).

In unserem Beispiel handelt es sich um eine umschriebene Fragestellung mit Identifikation der Patientencharakteristika (Patienten mit HIV-Infektion und einer latenten Tuberkulose-Infektion), Beschreibung der Intervention (Isoniazid-Prophylaxe) und Quantifizierung des therapeutischen Nutzens der Isoniazid-Prophylaxe.

Sind die Ein- und Ausschlusskriterien, nach denen die Studien berücksichtigt wurden, detailliert aufgeführt?

Die Ein- und Ausschlusskriterien, nach denen die Studien selektioniert wurden, müssen angegeben sein. Diese Kriterien sollten die Art der untersuchten Patienten und die therapeutischen Maßnahmen sowie die Art der Studienendpunkte beinhalten. Unterschiede bezüglich Studienpatienten (Alter, Schweregrad der Erkrankung), Medikamentenexposition und Expositionsdauer sowie die in den Studien primär berücksichtigten Endpunkte können zu unterschiedlichen Ergebnissen von Meta-Analysen führen, die die gleiche Studienfrage untersuchen [10]. Die methodologischen Standards, gemäß denen Studien selektioniert wurden, müssen erwähnt sein. Geht es um eine therapeutische Maßnahme, sollten nur randomisierte kontrollierte Studien eingeschlossen werden, da Ergebnisse von nicht randomisierten Studien wie z.B. Kohortenstudien durch Selektionsphänomene der Patientenpopulationen verfälscht sein können (Selektions-Bias). Die berücksichtigten Studien sollten nach dem Intention-to-treat-Prinzip analysiert (vgl. Kap. 12) worden sein.

Bei der Meta-Analyse unseres Beispiels handelt es sich um eine gepoolte Analyse von sieben randomisierten Studien mit Intention-to-treat-Analyse. Die evaluierten Studienendpunkte sind Tuberkulose, Tod sowie Nebenwirkungen der Therapie mit Isoniazid. Die in den Studien verwendeten Definitionen einer Tuberkulose sind allerdings nicht einheitlich. Nur eine der sieben Studi-

en verlangte einen Kulturnachweis für die Diagnose einer Tuberkulose. Prognostisch wichtige Patientenmerkmale wie Tuberkulinstatus und das Erkrankungsrisiko in den Placebogruppen sind aufgeführt. Hingegen fehlten in der Mehrzahl der Studien Angaben zu prognostischen Markern (CD4-Zellzahl). Die Qualität der eingeschlossenen Studien (Methodik der Randomisierung, Verblindung von Studienärzten und Patienten, Ausmaß des Follow-up von Studienpatienten) wurde ausgewertet.

13.3.2 Wie wahrscheinlich ist es, dass relevante Studien nicht berücksichtigt wurden (Publication-Bias)?

Eine Übersichtsarbeit muss umfassend sein, deshalb sollten Angaben, wie die Literatur gesichtet und welche Suchmethoden verwendet wurden, vorhanden sein. Die berücksichtigten Datenbanken und verwendeten Suchbegriffe sollten angegeben sein. Gewisse Übersichtsarbeiten beschränken sich auf die Sichtung der englischen Literatur. Eine solche Einschränkung kann jedoch zu Verzerrungen führen. Es konnte gezeigt werden, dass beispielsweise deutschsprachige Autoren ihre Studien mit „positiven“ und statistisch signifikanten Ergebnissen häufiger in englisch- als in deutschsprachigen Zeitschriften publizieren [11]. Die Ergebnisse einer Meta-Analyse, welche nur Studien aus englischsprachigen Zeitschriften berücksichtigt, können eine Verfälschung des tatsächlichen Nutzens beispielsweise einer therapeutischen Intervention darstellen.

Welche Suchstrategien wurden verwendet, um unveröffentlichte Studien zu identifizieren?

Da Studien mit positiven Ergebnissen häufiger publiziert werden als solche mit negativen Ergebnissen, kann der Nicht-Einbezug von negativen Studien zu einer Verzer-

rung des Ergebnisses einer Meta-Analyse führen (Publication-Bias). Aus diesem Grund wurde das internationale Register für randomisierte Studien eingeführt und Herausgeber namhafter Zeitschriften verlangen nun die Registrierung von Studien als Voraussetzung zur Publikation [12]. Es ist deshalb wichtig, dass Autoren von Übersichtsarbeiten Angaben machen, welche Anstrengungen sie unternommen haben, um nicht publizierte Studien zu identifizieren. Dies beinhaltet die Durchsicht von Abstracts (z.B. von Kongressberichten etc.) und den direkten Kontakt mit den Autoren der Primärstudien.

In unserem Beispiel sind die Suchstrategie und die verwendeten Suchbegriffe im Detail aufgeführt. Die Autoren der Primärstudien wurden für fehlende Informationen kontaktiert und WHO-Experten des Forschungsgebietes wurden befragt. Es ist somit eher unwahrscheinlich, dass die Suche unvollständig ist.

Wurde die Validität der eingeschlossenen Studien überprüft?

Auch wenn eine Meta-Analyse nur randomisierte kontrollierte Studien einschließt und somit ein gewisser Schutz vor Verzerrungseffekten (Selektions-Bias) gegenüber einer Meta-Analyse von z.B. Kohortenstudien besteht, ist es wichtig zu wissen, ob die Qualität der eingeschlossenen Studien gut war. Leider garantiert die Publikation einer Arbeit in einer Zeitschrift mit externer Beurteilung („Peer-Review“) die Validität von Studienergebnissen nicht. Deshalb müssen Angaben zur Validität der eingeschlossenen Studien nachvollziehbar sein. Empirische Studien haben gezeigt, dass sowohl randomisierte kontrollierte Studien ohne Verblindung von Patienten, Studienärzten als auch Studien ohne verblindete Endpunktbestimmung und Studien ohne verdeckte Randomisierung von Studienpatienten statistisch signifikant höhere Effektgrößen zeigen als Studien mit adäquater Verblindung. Der Ein-

schluss von methodologisch ungenügenden Studien in Meta-Analysen kann also zu einer Überschätzung des tatsächlichen Behandlungsnutzens führen [13; 14].

In unserem Beispiel haben die Autoren die Studienqualität mit einem Score [15] beurteilt und in einer Sensitivitätsanalyse den Effekt der unterschiedlichen Studienqualität auf die Gesamtergebnisse zusätzlich überprüft. Empirische Studien zeigen jedoch, dass die Studienqualität in Meta-Analysen besser anhand der oben erwähnten Einzelkomponenten und nicht mittels Qualitäts-Scores überprüft wird [16]. Allerdings ist die statistische Kraft, in einer Sensitivitätsanalyse relevante Unterschiede von Effektgrößen zwischen Studien mit unterschiedlichen Qualitätsmerkmalen zu finden, oft zu gering. Dies war auch in unserem Beispiel der Fall.

Ist die Art, wie die Studien überprüft wurden, reproduzierbar?

Studienselektion und Evaluation der Validität von Studiendesigns beinhalten wichtige Entscheidungen im Reviewprozess, die systematischen (*bias*) oder nicht systematischen Fehlern unterworfen sind. Der Einbezug von zwei oder mehreren unabhängigen Reviewern in diesen Entscheidungsprozess kann bei guter Übereinstimmung dem Leser Auskunft über die Zuverlässigkeit einer Übersichtsarbeit geben. Das Maß der Übereinstimmung kann mit einer „kappa-Statistik“ angegeben werden. Werte von kappa über 0,6 können als befriedigend betrachtet werden (die Streubreite der kappa-Statistik bewegt sich definitionsgemäß zwischen 0 und 1).

In unserem Beispiel haben zwei unabhängige Reviewer die relevanten Studien selektioniert und die Studienqualität überprüft. Ihre Übereinstimmung war sehr gut (kappa von 1 für die Studienselektion respektive 0,8 für die Studienqualität).

13.3.3 Wie lauten die Studienergebnisse?

Was sind die Gesamtergebnisse der Meta-Analyse?

Die Ergebnisse mehrerer Studien müssen in irgendeiner Form zusammengefasst werden. Das Auflisten und Vergleichen von positiven und negativen Ergebnissen in einer rein deskriptiven Zusammenfassung ist aus verschiedenen Gründen unbefriedigend. Studien erhalten unabhängig von ihrer Größe und Qualität ein gleiches Gewicht. Es besteht die Tendenz, Studien mit statistisch nicht signifikanten Ergebnissen, die klinisch relevant sein können, unterzugewichten. Die Schlussfolgerungen können somit bei gleicher Datenlage divergieren.

Eine Meta-Analyse bietet deshalb den Vorteil, die Ergebnisse einzelner Studien mittels einer quantifizierbaren statistischen Methode zusammenzufassen. Hierbei erhalten die Ergebnisse von größeren Studien ein größeres Gewicht als solche von kleinen Studien. Das Gesamtergebnis stellt somit ein gewichtetes Ergebnis aller individuellen Studien dar. Die Ergebnisse der gepoolten Analyse werden typischerweise als Odds-Ratios oder Risk-Ratios angegeben. Etwas vereinfachend gesagt kann der Kliniker Odds-Ratios oder Risk-Ratios gleich interpretieren wie die Angabe eines Relativen Risikos oder einer Relativen Risiko-Reduktion aus einer Einzelstudie. Die Ergebnisse einer Meta-Analyse sollten in gleicher Weise analysiert werden wie solche einer individuellen Behandlungsstudie, d.h. es sollten sowohl die Absolute als auch die Relative Risiko-Reduktion betrachtet werden.

In unserem Beispiel wurden sieben Studien mit insgesamt 2367 Patienten mit einer Isoniazid-Prophylaxe und 2162 Patienten mit Placebo evaluiert. Die Risk-Ratio einer Tuberkulose bei Prophylaxe mit Isoniazid war 0,58 (95%-Konfidenzintervall 0,43 bis 0,80). Hingegen zeigte sich unter Isoniazid-Prophylaxe kein Überlebensvorteil. Tenden-

ziell hatten Patienten mit Isoniazid mehr Nebenwirkungen (Risk-Ratio 1,36, 95%-Konfidenzintervall 1,00 bis 1,86).

Wie exakt sind die Ergebnisse (Vertrauensintervalle)?

Das Gesamtergebnis der Risiko-Reduktion zeigt den durchschnittlichen Effekt z.B. einer Behandlung gemittelt über alle Studien. Dieses Ergebnis ist ein Schätzwert, um den sich ein Vertrauensbereich (z.B. 95%-Konfidenzintervall) ermitteln lässt, der die gefundene Wirkgröße umschließt (vgl. Kap. 12).

In unserem Beispiel findet sich für den Hauptendpunkt (Inzidenz der Tuberkulose) sowohl in der Gesamtanalyse als auch in der Untergruppenanalyse bei Patienten mit positivem Tuberkulintest eine hohe Wirkgröße mit niedriger Risk-Ratio respektive hoher Risiko-Reduktion und einem relativ engen Vertrauensintervall – d.h. die statistische Kraft und die Zuverlässigkeit der gefundenen Wirkgröße ist hoch.

Sind die Studienergebnisse in den einzelnen Studien vergleichbar?

Trotz strikter Einschlusskriterien finden sich bei systematischen Übersichtsarbeiten Unterschiede in den berücksichtigten Studien, die auf unterschiedliche Patientencharakteristika und Interventionsdauer oder -intensität zurückgeführt werden können. Man sollte deshalb kritisch beurteilen, ob Unterschiede zwischen Studien deren Kombination grundsätzlich zulassen, oder anders ausgedrückt, ob allen Studien ein gleichartiger Behandlungseffekt zugrunde liegt. Diese Hypothese kann in Meta-Analysen mittels des so genannten Heterogenitätstests untersucht werden. Ist der Heterogenitätstest statistisch signifikant, dann bestehen zwischen den eingeschlossenen Studien unterschiedliche Effektgrößen. Bei Heterogenitätsnachweis muss das Gesamtergebnis und das Vertrauensintervall mit Vorsicht interpretiert wer-

den. Dann sollte besonderes Augenmerk darauf gerichtet werden, wie die Autoren die gefundene Heterogenität erklären. Leider schließt ein statistisch nicht signifikanter Heterogenitätstest das Vorhandensein von Heterogenität nicht aus, da der Test insbesondere bei Meta-Analysen von kleinen Studien eine geringe statistische Kraft hat und zudem von der Anzahl der eingeschlossenen Studien abhängt. Ein besseres Maß zur Beurteilung von Heterogenität ist das Ausmaß an Konsistenz I^2 , das durch die gesamte Heterogenität zwischen Studien und nicht durch den Zufall erklärt wird [17]. Der Konsistenzwert wird in Prozent angegeben und berechnet sich aus dem Q-Wert des Cochrane-Heterogenitätstests und der Anzahl der Studien (Freiheitsgrad df) $I^2 = 100\% \times (Q-df)/Q$. Ein Wert von 0% zeigt keine Heterogenität an, Werte zwischen 25%, 50% und 75% können als Maß von niedriger, moderater und hoher Heterogenität interpretiert werden.

Methodisch gute Meta-Analysen untersuchen in einer Sensitivitätsanalyse zusätzlich die Robustheit der postulierten, allen eingeschlossenen Studien unterliegenden Gesamtwirkgröße anhand von a priori festgelegten Parametern. Diese Parameter können z.B. unterschiedliche Qualitätsmerkmale der eingeschlossenen Studien oder Patientencharakteristika wie Alter, Risikofaktoren oder Krankheitsstadium sein.

In der vorliegenden Meta-Analyse fanden die Autoren für sämtliche untersuchten Endpunkte keine Heterogenität (d.h. der Heterogenitätstest war statistisch nicht signifikant). Hingegen war die Beobachtungsdauer in den Studien sehr unterschiedlich und eine Studie mit längerem Follow-up zeigte eine verminderte Effektivität der Isoniazid-Prophylaxe. Die Autoren erwähnen diesen Punkt in der Diskussion und haben zudem eine Sensitivitätsanalyse durchgeführt. Hierbei zeigte sich, dass der Nutzen der Isoniazid-Prophylaxe zur Tuberkuloseprävention sich auf Patienten mit positivem Hauttest (Induration >5mm)

beschränkte (Risk-Ratio 0,40, 95%-Konfidenzintervall 0,24 bis 0,65) und bei negativem Hauttest nicht nachweisbar war (Risk-Ratio 0,84, 95%-Konfidenzintervall 0,53 bis 1,30). Der Unterschied dieser beiden Summenergebnisse war statistisch signifikant ($p < 0,05$). Für die Mortalität ergab sich kein Unterschied. Ebenso fanden sich in Studien mit schlechteren gegenüber solchen mit besseren Qualitäts-Scores keine unterschiedlichen Effektgrößen.

13.3.4 Sind die Ergebnisse für die Behandlung meiner Patienten nützlich?

Können die Ergebnisse auf den/die von mir behandelten Patienten/in übertragen werden?

Ein Vorteil einer Meta-Analyse ist, dass die Ergebnisse sich aus einem breiten Spektrum von Patienten rekrutieren. Falls die Ergebnisse in allen Studien konsistent sind, sind sie auch auf eine breite Patientengruppe anwendbar. Trotzdem können Studienergebnisse häufig nur bedingt auf eine konkrete Patientensituation angewendet werden. Der zu behandelnde Patient ist z.B. älter oder weist eine höhere Komorbidität und somit schlechtere Prognose auf als die in den Studien untersuchten Personen. Solche Probleme werfen die Frage nach Untergruppenanalysen auf [18]. Generell sollten Schlussfolgerungen, die sich auf Vergleiche von Patienten-Subpopulationen zwischen verschiedenen Studien stützen, vorsichtig interpretiert werden. Beispielsweise fand eine Meta-Analyse über den Nutzen der β -Blocker-Therapie nach Herzinfarkt einen klinisch und statistisch signifikanten Unterschied zwischen Studien mit Präparaten mit beziehungsweise ohne intrinsischen sympathomimetischen Effekt [19]. Dieses Ergebnis bestätigte sich nach dem Einschluss von zwei später publizierten Studien jedoch nicht mehr [20]. Wei-

tere Kriterien, die das Ergebnis einer Untergruppenanalyse glaubhaft machen, sind ein großer Unterschied der Wirkgröße zwischen zwei Untergruppen mit einem signifikanten p-Wert für die Differenz des unterschiedlichen Behandlungseffektes. Hypothesen zu Untergruppenanalysen sollten a priori festgelegt und die Anzahl der Untergruppenanalysen begrenzt sein. Mit zunehmender Zahl von Untergruppenvergleichen steigt die Anzahl rein zufallsbedingter Unterschiede. Die Konsistenz der gefundenen Unterschiede gemittelt über alle Studien und eine plausible theoretische Begründbarkeit (Biologic Plausibility) unterstreichen die Wahrscheinlichkeit, dass die Ergebnisse der Untergruppenanalyse richtig und nicht zufallsbedingt sind.

In unserem Beispiel war die Anzahl der Untergruppenanalysen begrenzt. Zudem handelt es sich bei der Untergruppenanalyse gemäß Tuberkulintest-Status um Vergleiche von Merkmalen innerhalb von Primärstudien und nicht um Vergleiche von unterschiedlichen Merkmalen zwischen den Primärstudien. Dies und die Tatsache, dass die gefundenen Unterschiede statistisch signifikant waren, sind weitere Argumente dafür, dass die gefundenen Unterschiede zwischen den Untergruppen glaubhaft sind.

Der zu erwartende Behandlungsnutzen der Patientin kann in unserem Beispiel mittels einer einfachen Berechnung besser abgeschätzt werden [21] (s. Tab. 13.2). In unserer Meta-Analyse finden sich Informationen über das Erkrankungsrisiko von Patienten in den Kontrollgruppen mit positivem Tuberkulintest. In den in Afrika durchgeführten Studien betrug das Erkrankungsrisiko einer Tuberkulose zwischen 3% und 9% pro Jahr. Zum Vergleich: In Nordamerika und Mitteleuropa schwankt das Erkrankungsrisiko für Tuberkulose bei HIV-positiven Patienten zwischen 0,5% und 2,5%. Die in der Meta-Analyse gefundene Relative Risiko-Reduktion für eine Tuberkulose bei positivem Tuberkulin-

Tab. 13.2: Individueller Behandlungsnutzen (Number needed to treat) der Tuberkulose-Prophylaxe bei positivem intrakutanem Tuberkulintest (> 5 mm Erhabenheit) bei Patientinnen und Patienten mit einer HIV-Infektion in Abhängigkeit des Erkrankungsrisikos einer Tuberkulose

Erkrankungsrisiko an Tuberkulose in Prozent bei HIV Infektion und positivem Tuberkulintest	Tuberkulose Risiko, wenn unbehandelt (pro Jahr)	Tuberkulose Risiko, wenn behandelt (pro Jahr)	Relative Risiko-Reduktion	Absolute Risiko-Reduktion	Anzahl zu Behandelnder pro Jahr, um eine Tuberkulose zu verhüten
	B	U	U-B/U	U - B	1/(U-B)
Hohes Risiko (Afrika)	9%	3,6%	60%	5,4%	19
Mittelhohes Risiko (Afrika, i.v. Drogenbenützer Südeuropa)	3%	1,2%	60%	1,8%	56
Niedriges Risiko (Mitteleuropa)	1%	0,4%	60%	0,6%	167

Erkrankungsrisiko an Tuberkulose in Prozent bei HIV Infektion und positivem Tuberkulintest nach [2]

test liegt bei 60% (95%-Konfidenzintervall 35% bis 76%). Bei einem angenommenen Erkrankungsrisiko von 3% reduziert sich das Erkrankungsrisiko um 60% auf 1,2% (Berechnung der Risikodifferenz: $0,03 \times 0,6 = 0,018$). Die Anzahl der zu behandelnden Patienten (Number needed to treat/NNT), um eine Tuberkulose zu verhüten ist $1 / \text{absolute Risikodifferenz}$ ($1/0,018 = 56$). Bei einem angenommenen Erkrankungsrisiko von 9% beträgt die NNT 19. Diesem Behandlungsnutzen steht laut Meta-Analyse eine Risk-Ratio von 1,8 einer Isoniazid-induzierten Hepatopathie (definiert als Transaminasenanstieg über die zweifache obere Normgrenze) gegenüber.

Wurden alle relevanten klinischen Endpunkte in der Analyse berücksichtigt?

Systematische Übersichtsarbeiten mit einer fokussierten Fragestellung können bessere und valide Informationen über den Nutzen einer Behandlungsmaßnahme bieten. Entscheidend ist jedoch, ob die klinisch relevanten Endpunkte in die Analyse eingeschlossen werden. Eine systematische Übersichtsarbeit, die beispielsweise den Nutzen einer Substanz zur Osteoporose-Prävention

anhand des Endpunktes Knochendichte untersucht, gibt keinen Aufschluss über die wichtigere Frage, ob Patientinnen unter der Prophylaxe weniger Frakturen erleiden.

In der vorliegenden Meta-Analyse wurden sowohl „harte“ Endpunkte wie die Inzidenz der Tuberkulose und die Gesamtsterblichkeit als auch die Nebenwirkungen der Prophylaxe berücksichtigt. Hierbei gilt es jedoch zu berücksichtigen, dass der Nachweis der Tuberkulose nicht in allen Studien mittels Kultur erfolgte.

Überwiegen die positiven Auswirkungen die schädlichen Auswirkungen und Kosten der Maßnahme?

Die klinische Entscheidungsfindung beinhaltet das Abwägen der Risiken und des Nutzens einer therapeutischen Maßnahme unter besonderer Berücksichtigung der jeweils individuellen Patientensituation. Eine valide systematische Übersichtsarbeit kann eine gute Informationsgrundlage zur klinischen Entscheidungsfindung sein, die der jeweiligen individuellen Patientensituation angepasst werden muss.

13.4 Auflösung des klinischen Szenarios

Unsere Patientin stammt aus einem afrikanischen Land mit einer hohen Prävalenz der Tuberkulose, das demjenigen von HIV-infizierten Patienten in den Primärstudien entspricht. Bei einem Erkrankungsrisiko von 3% bis 9% wäre der zu erwartende Therapienutzen zwischen 19 und 56 zu Behandelnde pro verhütetem Tuberkulosefall, was Sie als günstig erachten. Etwas anders muss der Behandlungsnutzen beurteilt werden bei einem positivem Tuberkulintest bei einem HIV-positiven Patienten mit Herkunft aus einem Land mit niedriger Tuberkuloseprävalenz. Falls ein kürzlicher Kontakt mit einem Tuberkulose-Erkrankten ausgeschlossen werden kann, ist in letzterer Situation der zu erwartende Behandlungsnutzen wesentlich geringer.

Die zu erwartenden Nebenwirkungen, insbesondere das Risiko eines Isoniazid-bedingten Transaminasenanstiegs, halten Sie für vertretbar. Die Meta-Analyse ergab keine klaren Anhaltspunkte eines erhöhten Risikos der Tuberkulose-Prophylaxe bei HIV-Infizierten. Die in die Meta-Analyse eingeschlossenen Studien stammten mit einer Ausnahme aus Entwicklungsländern, wo keine antiretrovirale Therapie zur Verfügung stand. Es bestehen deshalb keine Daten über den Nutzen und die Risiken der Isoniazid-Prophylaxe bei gleichzeitiger antiretroviraler Kombinationstherapie.

Sie empfehlen der Patientin eine 9-monatige Isoniazid-Prophylaxe (300mg pro Tag) und die Patientin willigt nach erfolgter Information ein. Gleichzeitig bemühen Sie sich um die Kostenübernahme für eine antiretrovirale Kombinationstherapie.

Literatur

[1] Antman EM et al., A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA. (1992), 268, 240–248

- [2] Ravnskov U, Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome. BMJ. (1992), 305, 15–19
- [3] Lau J et al., Cumulative meta-analysis of therapeutic trials for myocardial infarction. N Engl J Med. (1992), 327, 248–254
- [4] Ioannidis JP, Cappelleri JC, Lau J, Issues in comparisons between meta-analyses and large trials. JAMA. (1998), 279, 1089–1093
- [5] Yusuf S, Teo K, Woods K, Intravenous magnesium in acute myocardial infarction. An effective, safe, simple, and inexpensive intervention. Circulation. (1993), 87, 2043–2046
- [6] ISIS-4: a randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. ISIS-4 (Fourth International Study of Infarct Survival) Collaborative Group. Lancet. (1995), 345, 669–685
- [7] Egger M, Smith GD, Misleading meta-analysis. BMJ. (1995), 310, 752–754
- [8] Oxman AD, Cook DJ, Guyatt GH, Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. JAMA. (1994), 272, 1367–1371
- [9] Bucher HC et al., Isoniazid prophylaxis for tuberculosis in HIV infection: a meta-analysis of randomized controlled trials. AIDS. (1999), 13, 501–507
- [10] Chalmers TC et al., Meta-analysis of clinical trials as a scientific discipline. II: Replicate variability and comparison of studies that agree and disagree. Stat Med. (1987), 6, 733–744
- [11] Egger M et al., Language bias in randomised controlled trials published in English and German. Lancet. (1997), 350, 326–329
- [12] DeAngelis CD et al., Is This Clinical Trial Fully Registered? A Statement From the International Committee of Medical Journal Editors. JAMA. (2005) 293, 2927–2929
- [13] Califf RM et al., A randomized controlled trial of eprostamol therapy for severe congestive heart failure: The Flolan International Randomized Survival Trial (FIRST). Am Heart J. (1997), 134, 44–54
- [14] Moher D et al., Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet. (1998), 352, 609–613

- [15] Jadad AR et al., Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. (1996), 17, 1–12
- [16] Juni P, et al., The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*. (1999), 282, 1054–1060
- [17] Higgins JP et al., Measuring inconsistency in meta-analyses. *BMJ*. (2003), 327, 557–60
- [18] Oxman AD, Guyatt GH, A consumer's guide to subgroup analyses. *Ann Intern Med*. (1992), 116, 78–84
- [19] Yusuf S et al., Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovasc Dis*. (1985), 27, 335–371
- [20] Peto R, Why do we need systematic overviews of randomized trials? *Stat Med*. (1987), 6, 233–244
- [21] Marx A, Bucher HC, Numbers needed to treat derived from meta-analysis: a word of caution. *ACP J Club*. (2003), 138(2), A11–2

14 Kritische Bewertung von qualitativen Studien

Thorsten Meyer

Eine Warnung vorweg: Die Auseinandersetzung mit der Bedeutung qualitativer Studien für die Evidenzbasierte Medizin ist ohne Risiken und Nebenwirkungen nicht zu leisten. Auf der einen Seite ist sie mit einer besonderen Herausforderung verbunden, da qualitative Studien – zumindest in einigen ihrer „Reinformen“ – auf einem grundsätzlich anderen **Forschungsparadigma** als Studien aus dem epidemiologischen, naturwissenschaftlich orientierten Bereich basieren, d.h. diese Auseinandersetzung erfordert die Bereitschaft, grundlegende Voraussetzungen der so genannten quantitativen Forschung in Frage zu stellen. Das ist mühsam. Erschwerend kommt hinzu, dass es auf Seiten der qualitativen Forschung keine **einheitlichen** Vorstellungen über das eigene Forschungsverständnis und somit nicht einmal über ihre Definition gibt. Auf der positiven Seite dieser Mühen steht die Möglichkeit, nicht nur etwas über eine wesentliche Forschungstradition der Human- und Geisteswissenschaften, sondern auch über die Grundannahmen des eigenen naturwissenschaftlich geprägten Forschungsverständnisses zu erfahren – und sie damit möglicherweise in einem kritischeren Licht zu betrachten.

Qualitative Studien stellen für die EbM eine neue Herausforderung dar, was sich z.B. daran zeigt, dass sie in der ersten Auflage dieses Lehrbuchs noch nicht in einem eigenen Kapitel gewürdigt wurden. Die Protagonisten der EbM waren mit den Methoden und Techniken der klinischen Epidemiologie vertraut, die auch heute einen herausragenden Stellenwert in der Anwendung der EbM aufweisen. Die klinische Epidemiologie weist eine

deutliche Orientierung an der naturwissenschaftlichen Forschung auf (z.B. hohe Wertschätzung der randomisierten kontrollierten Studie in Analogie zum naturwissenschaftlichen Experiment), ihre Ergebnisse stellen primär Quantifizierungen von Variablen über Patientengruppen dar. Ausgangspunkt der EbM ist jedoch nicht die Methode, sondern vielmehr das klinische Problem. Hier zeigte sich, dass bestimmte klinisch relevante Fragestellungen nicht mit den Methoden der klinischen Epidemiologie zu beantworten sind [1; 2], z.B.: Warum nehmen Patienten mit Hypertonie die von ihrem Arzt verschriebenen Medikamente nicht ein? Warum fangen Patienten nach einem lebensbedrohlichen KHK-Ereignis wieder an zu rauchen? Wie kann die unterschiedliche Verschreibungspraxis von konventionellen und atypischen Neuroleptika bei niedergelassenen Psychiatern erklärt werden? Welche Umstände hindern Ärzte daran, EbM in ihren Behandlungsalltag zu integrieren? Zum anderen gibt es Fragestellungen, zu deren Beantwortung sowohl Methoden der klinischen Epidemiologie als auch der qualitativen Forschung herangezogen werden können und sollten, z.B.: Wie ist der Erfolg einer psychoedukativen Maßnahme aus Sicht der betroffenenen, an Schizophrenie Erkrankten zu bewerten? Welche Faktoren bedingen die Inanspruchnahme medizinischer Leistungen von Personen mit chronischen Rückenschmerzen?

Bevor genauer darauf eingegangen wird, was qualitative Studien kenn- bzw. auszeichnet, soll ein klinischer Fall den Ausgangspunkt dieser Ausführungen setzen.

14.1 Klinisches Szenario

Wieder einmal sitzt Ihnen in Ihrer Hausarztpraxis ein Patient gegenüber, der beteuert, Ihre Ratschläge ernst genommen zu haben. Bei Herrn Dieter M. wurde vor mittlerweile fünf Jahren ein Diabetes mellitus Typ 2 diagnostiziert. Er ist 52 Jahre alt und arbeitet in der Personalverwaltung eines größeren Bauunternehmens. Er habe regelmäßig seine Medikamente eingenommen, in letzter Zeit weniger fetthaltiges Essen zu sich genommen, schon seit langem nicht mehr zur Zigarette gegriffen und er mache mittlerweile Sonntagsausflüge mit dem Fahrrad. Seine Blutwerte sprechen allerdings eine andere Sprache: sein HbA1c-Wert ist mit 8,6 weiterhin bedenklich hoch, mit einer über die Jahre eher steigenden Tendenz. Ein bekannter Hypertonus ist trotz Beta-blocker nur auf einem Niveau von 155/85 mmHg eingestellt, die Blutfette bewegen sich oberhalb der Norm (Serumcholesterin 8,9 mmol/l; HDL 1,1 mmol/l, LDL 6,9 mmol/l.). Auch sein Body-Mass-Index (BMI) von 28,7 (91 kg Gewicht bei 1,78 m Körpergröße) hat eher eine steigende Tendenz. Eine mögliche Umstellung des Patienten auf eine intensiviertere Insulin-Therapie haben Sie vor drei Monaten bei Ihrem letzten Termin mit Herrn M. angesprochen, sie wurde zu diesem Zeitpunkt von Ihrem Patienten kategorisch abgelehnt.

Dieser Patient ist nicht der einzige, der in dieses Muster passt. Nicht wenige Ihrer aktuellen, noch von größeren Folgeschäden verschonten diabetischen Patienten erscheinen Ihnen irgendwie „unbeteiligt“ gegenüber ihrer Erkrankung. Dieser Eindruck steht im deutlichen Kontrast zu dem Leiden von Diabetikern, die mit den Spätfolgen ihrer Erkrankung zu kämpfen haben sowie deren Angehörigen, die ihre immobilen, sehbehinderten, multimorbiden Partner bzw. Eltern pflegen. Die Gründe für die von Ihnen angenommene ungenügende Bereitschaft, Ihre therapeutischen Vorgaben bzw. Ratschläge anzunehmen, können Sie nur vermuten. Zum einen bleibt im klinischen Alltag zu wenig zeitlicher Spielraum, um sich den Patienten intensiver zuwen-

den zu können. Zudem müssen Sie sich eingestehen, dass Ihre Expertise für eine eingehendere Exploration möglicher bedingender psychischer oder lebensweltlicher Umstände nur rudimentär ist.

Vor diesem Hintergrund beschließen Sie sich zu informieren, welche verschiedenen Erklärungen dafür vorliegen, warum Patienten mit Diabetes mellitus Typ 2 therapeutisch als notwendig erachtete Maßnahmen nur unzureichend befolgen. Dieses Wissen soll Ihnen helfen, mögliche Problemfelder auf Patientenseite direkt ansprechen zu können und angemessenere individuelle Lösungsstrategien mit den Betroffenen zu erarbeiten. Sie erhoffen sich dadurch wieder einen besseren Zugang zu Herrn M. und damit die Möglichkeit, ihn von der Notwendigkeit einer intensivierten Insulin-Therapie zu überzeugen.

14.2 Die Evidenz suchen

Wann immer es um medizinische Probleme geht, die menschliche Sichtweisen, Präferenzen, Bewertungen oder Entscheidungen zum Gegenstand haben, Verhaltensweisen, Interaktionen und soziale Sinnzusammenhänge umfassen, lohnt es sich, nach Studien aus dem Bereich der qualitativen Forschung zu suchen. Der erste Schritt der EbM, die Formulierung der klinischen Problemstellung, stellt dabei die geringste Hürde der Nutzung qualitativer Studien für die EbM dar. Schon bei der Suche nach passenden qualitativen Studien werden in der Regel andere Wege beschritten werden müssen als bei den anderen in diesem Buch behandelten Fragestellungen. Qualitative Primärstudien haben erst seit einigen Jahren Einzug in die klinische Literatur gehalten. Dazu hat die internationale pflegewissenschaftliche Literatur einen entscheidenden Beitrag geleistet [3]. Für die Medizin hat das British Medical Journal in den 1990er-Jahren eine zentrale Vorreiterrolle eingenommen, das beginnend mit einflussreichen methodischen Arbeiten [1; 4] seither

eine Vielzahl von qualitativen Studien veröffentlicht hat. Mittlerweile haben qualitative Studien auch Einzug in andere Zeitschriften, die primär einer naturwissenschaftlich ausgerichteten Wissenschaftlichkeit verpflichtet sind, gehalten, z.B. im Lancet [5; 6] oder im Journal of the American Medical Association [7–9]. Trotz dieses regelrechten Booms qualitativer Forschung in den Gesundheitswissenschaften sind wir noch weit davon entfernt, qualitative Studien in der breiten Masse gesundheitswissenschaftlicher bzw. klinischer Literatur zu finden, was vielschichtige Gründe hat. Qualitative Forschungsprojekte haben es deutlich schwerer, an Fördermittel zu kommen. Gute qualitative Studien sind aufwendig, insbesondere personal- und zeitintensiv. Daher liegt ein nicht unbedeutender Teil qualitativer Studien in Form von Qualifikationsarbeiten vor, d.h. in Form von Dissertationen, Magister- oder Diplomarbeiten o.Ä. Qualitative Studien zeichnen sich dadurch aus, dass ihre Methoden flexibel an den Untersuchungsgegenstand angepasst werden und sich auch im Laufe eines Projekts verändern können. Dadurch wird es umso wichtiger, eine angemessene Darstellung der Rationale ihrer Methoden vorzunehmen, was in gängigen Publikationsformaten nur unzureichend zu leisten ist. Zur Darstellung der Ergebnisse qualitativer Studien wird bei Interviewstudien in aller Regel

Textmaterial aus Gesprächen dargestellt; Einzelfälle werden kontrastiert, um zu verallgemeinerbaren Aussagen gelangen zu können – auch dieser Prozess ist auf wenige Seiten begrenzt kaum sinnvoll abzubilden. Nicht zuletzt trägt auch die Form der Darstellung, die sich nicht selten an literarischen Vorbildern orientiert, dazu bei, das Auffinden relevanter qualitativer Literatur zu einer Herausforderung eigener Art zu machen.

Insbesondere aus den seit einigen Jahren zu verzeichnenden Ansätzen zur Erstellung systematischer Übersichtsarbeiten über Inhalte qualitativer Studien [6; 10; 11] konnten Erfahrungen zur systematischen Literaturrecherche zusammengetragen werden. Tabelle 14.1 führt wichtige medizinische bzw. pflegewissenschaftliche Zeitschriften auf, die z.T. zahlreiche qualitative Studien beinhalten und gleichzeitig über PubMed/ MEDLINE leicht recherchierbar sind. Als weitere Datenbanken sei auf Sociological Abstracts, CINAHL, EMBASE, PsycINFO, PSYINDEX, SSCI und SCI verwiesen, die allerdings nicht frei öffentlich zugänglich sind. Einige Artikel mit gesundheitswissenschaftlichen Bezügen finden sich in der im Internet zugänglichen Zeitschrift Forum Qualitative Sozialforschung/Forum: Qualitative Social Research (www.qualitative-research.net/fqs/fqs.htm). Eine gründliche und umfassende Literaturrecherche, wie sie für systematische Über-

Tab. 14.1: In PubMed aufgeführte Zeitschriften mit medizinisch relevanten Originalbeiträgen (indikationsübergreifend) aus dem Bereich qualitativer Forschung (Auswahl)

- British Journal of General Practice
- British Medical Journal
- Disability & Rehabilitation
- Family Practice
- International Journal of Nursing Studies
- Journal of Advanced Nursing
- Journal of Evaluation in Clinical Practice
- Journal of Public Health Medicine
- Preventive Medicine
- Social Science and Medicine
- Sociology of Health and Illness

sichtsarbeiten notwendig ist, kann nur mit großem Aufwand betrieben werden. Viele Artikel werden aus o.g. Gründen in Buchform, in Form so genannter „grauer“ Literatur (u.a. Qualifikationsarbeiten, Projektberichte), oder in für die Medizin randständigen Zeitschriften zu finden sein. Da keine einheitlichen Begrifflichkeiten vorliegen, um mit Sicherheit qualitative Studien zu identifizieren, wird eine solche Analyse eine sehr hohe Rate von „falsch positiven“ Angaben in Kauf nehmen müssen, um eine angemessene Sensitivität der Recherche zu gewährleisten [11]. Für die praktische Suche von qualitativen Studien wurden erste Suchfilter von der Health Information Research Unit der McMaster Universität in Kanada untersucht (vgl. [12]).

Da im vorliegenden Fall eine hochspezifische Suchstrategie benötigt wird, wurde auf einen entsprechenden, von der McMaster Universität veröffentlichten Suchfilter für PubMed (<http://hiru.mcmaster.ca/hedges/Qualitative.htm>) zurückgegriffen. Neben den methodischen Suchbegriffen „**qualitative[ti/ab] OR themes[ti/ab]**“ wurden die inhaltlichen Begriffe „**Diabetes Mellitus, Type 2**“ [MeSH] in Kombination mit „**compliance**“ gewählt, sowie die Suche auf „**humans**“ und „**all adults**“ eingeschränkt (LIMITS). Unsere Suche erbrachte 14 Treffer. Einige Studien basieren auf der Perspektive der behandelnden Ärzte, was vor dem Hintergrund unserer Fragestellung sekundär ist. Mehrere Studien untersuchten Gruppen von Patienten, bei denen die Übertragbarkeit der Ergebnisse nicht gewährleistet ist (u.a. australische Ureinwohner) oder die nur einen Teil der von Ihnen als relevant erachteten Patientengruppe erfassen, z.B. mit Fokus auf Patientinnen. Drei neuere Arbeiten erscheinen geeignet: eine Studie aus Kroatien, bei der Patienten im Rahmen von Fokus-Gruppen zu ihren Wahrnehmungen ihrer Erkrankung und Behandlung befragt wurden [13], eine schwedische Studie, die über die Analyse der Interak-

tion von Arzt und Patienten relevante Wahrnehmungen der Patienten zu ermitteln suchte [14], sowie eine Studie aus Schottland, die auf der Basis von individuellen Interviews mit Patienten eine Analyse von Vorstellungen über Ursachen und Kontrolle des Diabetes darstellen [15]. Sie entscheiden sich dafür, in einem ersten Schritt die Studie von Parry et al. [15] für Ihre Frage heranzuziehen.

14.3 Die Evidenz bewerten – Kritische Bewertung einer qualitativen Studie

Qualitative Studien in der Medizin werden häufig mit ausführlichen Patienteninterviews assoziiert, die ihr Analogon in der ausführlichen klinischen Anamnese – und somit einer originär medizinischen Tätigkeit – aufweisen [16]. Gegenüber der klinischen Anamnese zeichnen sich verschiedene qualitative Ansätze durch ihre eigene explizite Systematik der Datengenerierung aus. Zudem ist der Einzelfall für den qualitativen Forscher primär Ausgangspunkt für eine induktive Theoriebildung, um über den Einzelfall hinausweisende Aussagen zu ermöglichen. Darüber hinaus beschränken sich qualitative Methoden nicht auf Interviews mit Betroffenen, vielmehr existiert eine in der medizinischen Literatur häufig übersehene Tradition der Verhaltensbeobachtung und Dokumentenanalyse [17; 18]. In Tabelle 14.2 finden sich beispielhaft verschiedene Ansätze und Methoden bzw. Methodologien qualitativer Forschung.

Vor der Einführung von Kriterien für die Bewertung qualitativer Studien sollen in einem kurzen Exkurs wesentliche Merkmale qualitativer Forschung vorgestellt werden. Qualitative und quantitative Forschung galten lange Zeit als zwei unversöhnliche Forschungstraditionen, die in aller Regel auch institutionell getrennt waren. Diese Trennung legt nahe, dass es sich um zwei grund-

Tab. 14.2: Beispiele für Inhalte und Methoden bzw. Methodologien qualitativer Studien

<ul style="list-style-type: none"> • Biographieforschung • Forschung über Krankheitsverläufe • Organisationsforschung • Geschlechterforschung • Qualitative Versorgungsforschung • Dokumenten- und Aktenanalyse 	<ul style="list-style-type: none"> • Leitfaden-Interview • Teilnehmende Beobachtung • Fokus-Gruppen • Kognitives Interview • Qualitative Inhaltsanalyse • Typenbildung 	<ul style="list-style-type: none"> • Grounded Theory • Action Research • Konversationsanalyse • Objektive Hermeneutik • Forschungsprogramm • Subjektive Theorien
---	--	--

sätzlich verschiedene und umschriebene Forschungsparadigmen handelte, worüber schon trefflich gestritten wurde [19; 20]. Der Begriff „Qualitative Forschung“ kennzeichnet ein sehr heterogenes Konglomerat unterschiedlichster Forschungsansätze, die sich kaum in einer einheitlichen Definition erfassen lassen. Vielmehr findet sich zur Charakterisierung qualitativer Forschung in aller Regel eine Gegenüberstellung mit Merkmalen der so genannten quantitativen Forschung [1], d.h. sie wird in der Regel in dieser Abgrenzung definiert und legitimiert.

Bedeutung sozialer Realitäten

Entgegen der naturwissenschaftlichen Vorstellung **einer** äußeren Realität, deren Elemente und Prinzipien erkannt und in Form von Gesetzmäßigkeiten formuliert werden können, geht die qualitative Forschung von verschiedenen Realitäten aus – eine Annahme, die insbesondere für die Beschreibung sozialer Phänomene notwendig erscheint. Der extremste Ansatz (Radikaler Konstruktivismus) verweist darauf, dass „die Wirklichkeit“ nur in unseren Köpfen konstruiert werden kann. Mit dem Fokus auf die soziale Umwelt machen Soziale Konstruktivistinnen darauf aufmerksam, dass die Art und Weise, wie wir Menschen unsere soziale Umwelt sehen und verstehen, soziokulturell geprägte Kategorien darstellt, die die Grundlage für unser Handeln und Verhalten bilden. Bspw. werden mit der Verwendung des Begriffs „Compliance“ zentrale Konzepte über die Arzt- und Patientenrolle verbunden, die wiederum die „sozialen Realitäten“ der beteiligten Personen prägen. In dem anfangs eingeführten

Fall wird unausgesprochen davon ausgegangen, dass die therapeutischen Ratschläge des Arztes den Maßstab darstellen. An diesem muss sich das Verhalten des Patienten messen, das entweder nur passend oder abweichend-defizitär sein kann. Qualitative Forscher sind in aller Regel sensibel für zugrunde liegende soziale (Macht-) Strukturen und Kategorien und die mit ihnen einhergehenden Wertsetzungen (z.B. paternalistisches Arzt-Patient-Verhältnis). Da die Beschreibung und Analyse der sozialen Realitäten wiederum nur von Forschern als Subjekte geleistet werden kann – und nicht als unvoreingenommene „objektive“ Beobachter – sind qualitative Forscher sehr darauf bedacht, ihren eigenen Standpunkt zu explizieren. Dies dient u.a. dazu, sich bewusst zu machen, an welchen Punkten im Forschungsprozess nicht die sozialen Wirklichkeiten der zu Untersuchenden, sondern vielmehr die eigenen Vorstellungen die Ergebnisse prägen. Die Subjektivität des Forschers – in der quantitativen Forschung als Bias oder unzureichende Objektivität verstanden – wird bei qualitativen Forschern nicht nur als unvermeidlich betrachtet, sondern in die verschiedenen Stadien des Forschungsprozesses integriert.

Betrachtung des Menschen in seinem natürlichen Lebensumfeld

Um die sozialen Wirklichkeiten und ihre Regeln zu verstehen, muss die Forschung möglichst nahe an diesen Realitäten ansetzen. Das Idealbild dieser Art von Forschung haben die Ethnologen/Kulturanthropologen geprägt, die sich den sozial-kulturellen Nor-

men und Werten von Volksgruppen durch Teilnahme am Alltagsleben, Erlernen ihrer Sprache und Rituale, d.h. einem Verstehen sozialer Phänomene von innen heraus angenähert haben. Als Gegenstück dieses Ansatzes wird die experimentelle Studie im Labor angesehen: künstlich von Versuchsleitern geschaffene Situationen, die für die Wirklichkeit der Betroffenen ohne Bedeutung sein können. Schon an diesem Punkt wird die Schwierigkeit der Abgrenzung von quantitativer und qualitativer Forschung deutlich. Quantitative Forschung kann genauso im natürlichen Umfeld stattfinden und Erhebungssituationen schaffen, die deutlich alltagsnäher sein können als Methoden aus dem qualitativen Kanon, wie z.B. Fokus-Gruppen.

Induktives vs. deduktives Schlussfolgern

Der wichtigste Ansatzpunkt qualitativer Forschung ist der Einzelfall. Allgemeinere Aussagen leiten sich immer aus der umfassenden Beschreibung und Analyse von Einzelfällen ab, somit stellt die **Induktion** einen zentralen Bestandteil des Schlussfolgerns dar. Die quantitative Forschung weist dagegen mit ihrer Fundierung im kritischen Rationalismus in der Tradition Karl Poppers nur deduktiven Schlüssen einen wissenschaftlichen Aussagewert zu; Hypothesen werden hier formuliert und mithilfe empirischer Daten an „der Wirklichkeit“ überprüft (bzw. angenommen falsifiziert oder nicht falsifiziert). Allerdings benötigt auch die quantitative Forschung Induktion, sie wird jedoch nur selten systematisch vorgenommen. Das Entwickeln von Theorien oder die Aufstellung von Hypothesen ist ohne induktives Denken nicht sinnvoll vorstellbar. Die qualitative Forschung hat durch ihre Systematisierung der Induktion einen wesentlichen Beitrag zur fundierten, expliziten Entwicklung von Theorien und Hypothesen geleistet. Das bedeutet allerdings nicht, dass auf deduktive Schlussfolgerungen verzichtet werden kann.

Vielmehr sind induktiv-deduktive Erkenntnisprozesse, die kennzeichnend für menschliche Wahrnehmung und Schlussfolgerung sind, substanziell in den Forschungsablauf integriert. Mehrfach durchlaufende induktiv-deduktive Erkenntnis Schleifen kennzeichnen die kontinuierliche Annäherung an die interessierenden Phänomene.

Wahrheitsbegriff

Einig werden sich die meisten qualitativen Forscher darin sein, dass sie nicht nach der Erkenntnis **einer objektiven** Wahrheit streben. Beschreibungen und Erklärungen der sozialen Realitäten sind immer kulturell und historisch gebunden. Damit erhält der Konsens zwischen Forschern bzw. zwischen Forschern und Beforschten eine zentrale Bedeutung für die Qualifizierung dessen, was als „wahr“ gelten kann. Die zentrale Frage der qualitativen Forschung ist nicht, ob ihre Erkenntnisse wahr sind oder nicht, sondern vielmehr, wie passend sie die gegebenen sozialen Wirklichkeiten beschreiben und erklären können. Dadurch, dass immer verschiedene passende Sichtweisen auf die zu beschreibenden Phänomene möglich sind, ist die vermeintliche Suche nach der „einen Wahrheit“ unsinnig. Damit soll nicht impliziert werden, dass epidemiologische oder andere quantifizierende Forschungsansätze nach der Erkenntnis der „einen Wahrheit“ streben; vielmehr werden auch in diesen Ansätzen mit der Abkehr vom Positivismus der unauflösbare Zusammenhang zwischen Wahrnehmung und Wirklichkeit bzw. die zeitlich-kulturelle Abhängigkeit des Wissens problematisiert.

Qualität qualitativer Forschung

Es gibt keine allgemein akzeptierten Qualitätskriterien zur Bewertung qualitativer Studien. Zentrale Problemfelder bei der Formulierung entsprechender Standards lassen sich aus den o.g. Merkmalen qualitativer Forschung ableiten. Während in der so genannt

ten quantitativen Forschung nach „einer Wahrheit“ gesucht wird (wobei man danach trachtet, das Ausmaß systematischer oder unsystematischer Verzerrung zu minimieren), ist der Ausgangspunkt qualitativer Forscher von mehreren „Wirklichkeiten“ bzw. zumindest verschiedenen wahrgenommenen Wirklichkeiten gekennzeichnet, die alle ihre eigene Berechtigung (und ihren eigenen Wahrheitsgehalt) haben. Auch die Interpretationen und Schlussfolgerungen der Forscher stellen prinzipiell nur eine von verschiedenen möglichen „Wirklichkeiten“ dar. Die Vielfalt qualitativer Ansätze mit ihren unterschiedlichen epistemologischen Ausgangspunkten (vgl. [18], Tabelle 2) beinhaltet auch eine Vielfalt methodischer Zugänge zur Forschungsfrage. Aus der Untersuchungsfrage soll argumentativ eine Methodologie abgeleitet werden, auf deren Grundlage eine Methodik entwickelt und begründet wird, die auf den Untersuchungsgegenstand zugeschnitten sein soll. Weiter kommt erschwerend für die Formulierung von Standards hinzu, dass sich die qualitative Forschung durch Offenheit gegenüber dem Untersuchungsgegenstand auszeichnet. Das bedeutet u.a., dass sich die ursprüngliche Fragestellung im Laufe der Untersuchung durchaus verändern kann, ebenso wie ihre Methoden [19]. Die Formulierung konkreter Qualitätskriterien, die sich im Rahmen einer Checkliste abarbeiten lassen, erscheint vor diesem Hintergrund sehr problematisch und ist entsprechend wiederholt grundsätzlich infrage gestellt bzw. kritisch hinterfragt worden [21]. Gleichzeitig muss sich jedes Forschungsprojekt unabhängig von seiner Forschungsmethodologie einer entsprechenden Qualitätsbeurteilung unterziehen, um die Förderung und Publikation der Ergebnisse zu ermöglichen. Die vorhandenen Ansätze der Formulierung von Qualitätskriterien stehen in diesem Spannungsfeld. Sie erfordern alle ein vertieftes Verständnis der Eigenarten qualitativer Forschung, und auch ihre Anwendung

erfordert eine argumentative Auseinandersetzung mit den dargestellten Inhalten qualitativer Studien, die nicht den Charakter einer Checkliste annehmen können. Murphy et al. [19] formulierten Qualitätskriterien im Rahmen einer umfassenden und vertieften Auseinandersetzung mit der Bedeutung qualitativer Studien für die Bewertung von Gesundheitstechnologien (HTA). Es gibt in den Gesundheitswissenschaften weitere Initiativen zur Bewertung qualitativer Studien, u.a. vom JAMA in ihrer Reihe „Users' guide to the medical literature“ [7; 8], vom Lancet [5; 22], aus pflegewissenschaftlicher Perspektive von Sandelowski & Barroso [23], im Rahmen der Cochrane Collaboration [24] sowie des britischen Critical Appraisal Skills Programm (CASP). Von Steinke [25] liegt eine ausführliche Darstellung und Analyse verschiedener Kriterien qualitativer Forschung aus der Perspektive qualitativ-empirischer Sozialforschung vor (vgl. Tab. 14.3). Eine fundierte Einführung in diese Kriterien würde weit über die Möglichkeiten dieses Kapitels hinausgehen. Sie sollen vielmehr das Gerüst für die Fragen darstellen, die an eine qualitative Studie gestellt werden sollten.

Die Formulierung von Kriterien qualitativer Forschung und Vorschläge zu ihrer konkreten Überprüfung sind begrifflich nicht immer klar voneinander getrennt worden. Bspw. findet sich in verschiedenen Publikationen der „Member Check“ bzw. die kommunikative Validierung als Qualitätskriterium, obschon sie kein übergreifendes Kriterium darstellen und in verschiedenen Ansätzen sogar kontraindiziert sein kann. Steinke [25] versteht Member Check entsprechend als eine mögliche Form, um das Kriterium der **empirischen Verankerung** zu überprüfen. Die Triangulation, d.h. die Kombination verschiedener Methodologien für die Untersuchung des gleichen Phänomens, wurde ebenfalls als „Kriterium“ qualitativer Forschung in der Literatur bezeichnet. Ihre verschiedenen Ausprägungen (Methoden-,

Tab. 14.3: Kriterien zur Bewertung qualitativer Studien und Wege zur Sicherung und Prüfung, in Anlehnung an Steinke [25] (gekürzt und modifiziert vom Autor)

- **Intersubjektive Nachvollziehbarkeit**
 - Dokumentation des Forschungsprozesses
 - Interpretation in Gruppen
 - Anwendung kodifizierter Verfahren (d.h. bereits entwickelter, aus Regeln bestehender Verfahren, z.B. Methode der Grounded Theory)
- **Indikation des Forschungsprozesses**
 - Indikation des qualitativen Vorgehens angesichts der Fragestellung
 - Indikation der Methodenwahl
 - Indikation von Transkriptionsregeln
 - Indikation der Samplingstrategie
 - Indikation der methodischen Einzelentscheidungen im Kontext der gesamten Untersuchung
- **Empirische Verankerung**
 - Verwendung kodifizierter Methoden (s.o.)
 - Hinreichende Textbelege
 - Reaktion auf Widersprüche, abweichende bzw. negative Fälle
 - Analytische Induktion als Methode der Theorieentwicklung
 - Ableitung und Überprüfung von Prognosen
 - Kommunikative Validierung (Member Check)
- **Limitationen (Beschreibung der Kontexte und Identifikation relevanter Bedingungen)**
 - Fallkontrastierung
 - Explizite Suche nach abweichenden, negativen und extremen Fällen
- **Kohärenz der Theorie**
 - Kohärenz der Theorie
 - Bearbeitung von Widersprüchen in den Daten und Interpretationen
- **Relevanz**
 - Relevanz der Fragestellung
 - Relevanz für Theorie
- **Reflektierte Subjektivität**
 - Integration von Selbstbeobachtung
 - Reflektion persönlicher Voraussetzungen
 - Vertrauensbeziehung zu den Beforschten
 - Reflektionen während des Feldeinstiegs
- **Multiperspektivität***
 - Triangulation von Methoden
 - Triangulation von Daten
 - Triangulation von Untersuchern
 - Triangulation von Theorien
- **Gerechter Umgang (Fair Dealing)***
 - Fairer Einbezug verschiedener am Untersuchungsphänomen beteiligter Perspektiven
 - Berücksichtigung von verschiedenen Statusgruppen/Machtunterschieden

* Diese Kriterien wurden der Liste von Steinke [25] vom Autor hinzugefügt, in Anlehnung an Murphy et al. [19]

Daten-, Untersucher- und Theorien-Triangulation) konkretisieren jedoch vielmehr das Kriterium der Multiperspektivität, d.h. der Annäherung an eine Fragestellung unter verschiedenen Sichtweisen.

Es wird deutlich, dass die Kriterien jeweils für verschiedene Abschnitte des Forschungsprozesses und damit auch für die Publikation der Studie von Bedeutung sind. Daher ist es für die pragmatische Beurteilung von qualitativen Studien in der EbM sinnvoll, auf der Grundlage der genannten Qualitätskriterien spezifische Fragen zu formulieren, um sie auf verschiedene Abschnitte einer Publikation anwenden zu können. Entsprechende Aufstellungen finden sich u.a. bei Mays & Pope [26], Malterud [5] sowie Sandelowski & Barroso [23]. Diese Strukturierung findet sich in gesundheitswissenschaftlichen Publikation häufig wieder; sie vernachlässigt allerdings den prozesshaften, zyklischen Charakter qualitativer Studien, in dem sich z.B. die Forschungsfrage im Verlauf der Untersuchung verändern kann, die Art der Stichprobenziehung von Ergebnissen abhängt, Interpretationen wieder erneute Datenerhebungen nach sich ziehen können etc. Im Folgenden soll eine kritische Würdigung der Studie von Parry et al. [15] mithilfe von an den Kriterien orientierten Fragen erfolgen.

14.4 Sind die Ergebnisse der Studie glaubwürdig?

Die im Folgenden aufgeführten Fragen dienen der Erfassung der oben aufgeführten Kriterien, bezogen auf die in einer Studie dargestellten Schritte. Sie sind weder umfassend noch im Einzelnen für alle verschiedenen Formen qualitativer Studien anwendbar. Es empfiehlt sich, auf der Grundlage der in Tabelle 14.3 aufgeführten Kriterien entsprechende Fragen zu formulieren, die zu der vorliegenden Studie passen. Eine Publikation

in einem Journal ist gleichzeitig in ihrem Umfang zu begrenzt, um alle o.g. Kriterien zufriedenstellend erfüllen zu können. Aus Sicht des Autors wären als Minimalforderung die Kriterien Intersubjektive Nachvollziehbarkeit, Empirische Verankerung, Reflektierte Subjektivität und Multiperspektivität positiv zu bewerten, um die Glaubwürdigkeit der Studienergebnisse zu stützen.

14.4.1 Forschungsfrage

Wurde dargestellt, welche Ausgangsfrage vorlag und ob bzw. wie sich die Forschungsfragen im Verlauf der Untersuchung geändert haben? Ist die Forschungsfrage spätestens am Ende der Untersuchung klar formuliert worden? Wurde die Forschungsfrage in Bezug zu vorhandenem Wissen bzw. vorliegenden Theorien gesetzt? Hat der Untersucher seine Vorannahmen über die Forschungsfrage expliziert? War er in der Lage, seine Vorannahmen zu relativieren?

Das übergeordnete Ziel der vorliegenden Studie bestand darin, „Diskurse“ zu explorieren, auf die Personen mit Diabetes mellitus Typ 2 in der Beschreibung ihrer Reaktionen auf die Erkrankung zurückgreifen, sowie die Implikationen der von ihnen übernommenen Standpunkte für den Umgang mit ihrer Erkrankung. Daraus sollten Empfehlungen für die zukünftige Versorgung von Diabetikern abgeleitet werden. Die Forschungsfrage wurde somit klar formuliert. In der Einleitung wird ein expliziter Bezug zu den Annahmen der Diskursanalyse vorgenommen, in einem kurzen Absatz auch inhaltlich zu einigen Diskursen, die in der Literatur beschrieben sind (Patienten in der Rolle als Opfer der Umstände; Patienten, die ihr eigenes Schicksal in den Händen haben). Über eine Anpassung der Fragestellung im Verlauf der Studie wurde nicht berichtet. Es finden sich keine Angaben dazu, wie konkret das Vorver-

ständnis der Autoren Eingang in die Formulierung der Forschungsfrage gefunden hat.

14.4.2 Untersuchungsdesign

War das angewendete Untersuchungsdesign in der Lage, die Forschungsfrage zu beantworten? Wurden alternative Untersuchungsdesigns diskutiert bzw. wurde die Wahl des Designs argumentativ sinnvoll gestützt? Wurden verschiedene relevante Perspektiven im Untersuchungsdesign integriert (Triangulation, Fair Dealing)?

Zur Identifikation von Diskursen bedarf es verbalen Materials von den beteiligten Protagonisten. In der vorliegenden Studie wurden neu erkrankte Personen mit Diabetes mellitus Typ 2 an drei Erhebungszeitpunkten im Abstand von jeweils sechs Monaten mittels eines leitfadengestützten Interviews befragt. Eine explizite, substanzielle Begründung für die Wahl der Mehrfacherhebungen wurde nicht gegeben. Aus der Darstellung der Ergebnisse und der Diskussion wurde jedoch deutlich, dass mithilfe dieses Designs Aussagen darüber getroffen werden konnten, wie konsistent die Befragten auf die identifizierten Diskurse zurückgegriffen haben, wie diese Diskurse im Zusammenhang mit ihrem Verhalten gegenüber professionellen Angeboten bzw. mit dem Umgang mit ihrer Erkrankung i.w.S. und letztlich auch mit Outcomes stehen.

Die Forschungsfrage fokussiert nicht primär auf die Identifikation gesellschaftlicher Diskurse, sondern auf Patienten, die auf diese Diskurse zurückgreifen. Aus diesem Grund wurde dem Fokus der Patientensichtweise durch die Methode leitfadengestützter Interviews Rechnung getragen. Von anderen an der Versorgung Beteiligten (z.B. Ärzte, Pflegepersonal, Angehörige) könnten nur indirekte Informationen erhalten werden. Andere Quellen, z.B. Dokumentationsmaterial, Beschwerdebriefe, Autobiografien, erscheinen

für dokumentierte Alltagserfahrungen im Umgang mit einer verbreiteten Erkrankung weniger geeignet. Eine kritische Diskussion möglicher alternativer methodischer Zugänge findet sich in der vorliegenden Publikation nicht.

14.4.3 Darstellung des Kontexts

Wurde der Kontext oder das Setting angemessen und umfassend dargestellt, sodass die Möglichkeit der Verallgemeinerbarkeit der Ergebnisse überprüft werden kann?

Es finden sich Hinweise zum sozioökonomischen Hintergrund der Befragten, zu ihrer Behandlung sowie zum äußeren Umfeld, einer Local Health Care Cooperative in Schottland zur Primärversorgung. Für die individuell dargestellten Fälle werden ebenfalls einige Hintergrundinformationen im Zuge der Ergebnisdarstellung gegeben. Insgesamt fällt die Darstellung des Kontextes eher knapp aus.

14.4.4 Stichprobengewinnung

Wurde eine theoretisch begründete Strategie zur Gewinnung der Stichprobe formuliert und eingehalten? Umfasste die Stichprobe den vollen Umfang möglicher Fälle oder Settings, sodass konzeptuelle Verallgemeinerungen möglich sind? Wurde Wert darauf gelegt, Fälle zu identifizieren, die den eigenen (vorläufigen) Vermutungen widersprechen können?

Über die Form und Begründung der Stichprobengewinnung wurden keine Angaben im Text gemacht. Entsprechend finden sich auch keine Hinweise auf Bemühungen, widersprüchliche Fälle zu identifizieren. Die Zahl der Interviewten ist zwar mit $n=40$ für eine qualitative Interviewstudie vergleichsweise hoch, allerdings ist die Frage nach einer möglichen unbeabsichtigten Selektion der Patienten nicht zu beantworten. Der Ver-

weis darauf, dass die Patienten in Geschlecht, Alter und sozioökonomischem Status variieren, legt nahe, dass sich die Autoren primär an dem Ideal einer repräsentativen Stichprobe orientiert haben, ohne dies jedoch zu spezifizieren.

14.4.5 Datenerhebung und -aufbereitung

Wurden ggf. angemessene technische Hilfsmittel verwendet, um die Akkuratheit der Daten zu gewährleisten (z.B. Tonbandaufzeichnung bei Interviews, Videodokumentation zur Analyse von Interaktionen)? Wurde ggf. auf eine systematische und den Daten angemessene Strategie der Datentranskription zurückgegriffen? Haben sich die Forscher ausreichend lange im Feld aufgehalten, insbesondere um diskordante Fälle entdecken zu können?

Es ist davon auszugehen, dass die Interviews aufgezeichnet wurden, da Interview-Transkripte vorliegen. Dabei wurde in normale Schriftsprache unter Beibehaltung grammatischer Fehler übertragen (im Gegensatz zu literarischen Umschriften, die auch Sprachfehler, Dialekte oder paraverbale Äußerungen beinhalten können), was aufgrund der primär inhaltlich-thematischen Bedeutungsebene der Texte angemessen ist. Aufgrund der Datenmenge (3 x 40 Interviews) wurde eine computer-gestützte Analysesoftware verwendet. Inwieweit die beteiligten Forscherinnen und Forscher sich zusätzlich zu den Interviews im Forschungsfeld aufgehalten haben bzw. mit diesem im Alltag vertraut sind, wurde nicht berichtet.

14.4.6 Datenanalyse und -interpretation

Ist die angewendete Form der Datenanalyse angemessen auf die Forschungsfrage und den Untersuchungsgegenstand abgestimmt? Wurden alternative Analyseme-

thoden diskutiert bzw. wurde die Wahl der Analysemethode argumentativ sinnvoll gestützt? Inwieweit konnten verschiedene Perspektiven von Betroffenen integriert werden? Wurde die Analyse so ausführlich dargestellt, dass sie von anderen Forschungsgruppen in ähnlicher Weise durchgeführt werden könnte?

Es sind nur wenige Hinweise vorhanden, wie die Datenanalyse vorgenommen wurde, wodurch eine Replikation erschwert wird. Die Autoren haben sich mit den Interviewinhalten durch mehrfaches Lesen der Transkripte vertraut gemacht. Aus den Daten sollten diskursive Konstruktionen identifiziert und mit biomedizinischen Diskursen in Beziehung gesetzt werden. Eine Diskussion verschiedener Analysemethoden ist nicht erfolgt, eine Einbettung der vorgenommenen Analyse in den Forschungsansatz der Diskursanalyse liegt nahe. Alle vier Autoren waren an der Analyse beteiligt, sodass Sichtweisen verschiedener Forscher in den Analyseprozess integriert waren.

In welchem Ausmaß wurden im Rahmen der Analyse Kategorien oder Konzepte entwickelt, die die zentralen Aussagen, Beobachtungen oder Prozesse angemessen erklären können? Liegen eindeutige Definitionen der entwickelten Kategorien oder Konzepte vor? Inwieweit kann die Iteration, d.h. die wiederholte Rückkopplung zwischen den Daten und den theoretischen Annahmen bzw. Konzepten, nachvollzogen werden? Wurde ein ausreichender Anteil an Daten aufgeführt, damit der Leser nachvollziehen kann, ob die Interpretationen aus den Daten gerechtfertigt sind? Haben die Forscher explizit nach „negativen Fällen“ gesucht bzw. diese in die Analyse einbezogen, d.h. Fälle, die den entwickelten Annahmen bzw. Konzepten widersprechen?

Die Autoren identifizierten zwei diskursive Konstruktionen (s.u.), die sie in Form von

Idealtypen bzw. Extremfällen ausführlich darstellten und ebenso Beispiele für Mischformen dieser Typen aufführten. Explizierte Definitionen dieser Konstruktionen erfolgten nicht, sie wurden dagegen ausführlich beschrieben und verschiedene relevante Dimensionen im Diskussionsteil aufgegriffen. So wie der Datenanalyseprozess undeutlich bleibt, ist auch die Iteration zwischen den Daten und den ermittelten Konzepten bzw. den eigenen Vorannahmen nicht expliziert worden. Mithilfe der aufgeführten Daten in Form von Zitaten wurde in ausreichendem Ausmaß die Interpretation in Bezug zu den identifizierten Konstruktionen nachvollziehbar. Eine explizite Suche nach „negativen Fällen“ findet sich nicht; vielmehr haben die Autoren deutlich gemacht, dass viele Patienten durch eine Mischung der beiden vorgestellten diskursiven Konstruktionen zu charakterisieren sind.

Wurden alternative plausible Erklärungsansätze für die Daten in Erwägung gezogen? Erfolgte eine Rückmeldung der Ergebnisse an die Befragten und wurden die Reaktionen auf diese Rückmeldung in die Datenanalyse integriert (Respondent Validation, Member Check)? Wurde darüber reflektiert, inwieweit die eigenen Vorannahmen die Datenanalyse beeinflusst haben?

Eine Rückmeldung der Ergebnisse an die Befragten erfolgte nicht. Da die Ergebnisse in Form von diskursiven Konstruktionen nur sehr bedingt auf den einzelnen Interviewten zu beziehen sind, stellt sich natürlich die Frage nach der Durchführbarkeit und Sinnhaftigkeit eines solchen Member Checks. Ebenso wenig finden sich Hinweise, ob darüber reflektiert wurde, inwieweit die eigenen Vorannahmen die Datenanalyse beeinflusst haben.

Zusammenfassend lässt sich festhalten: die Stärken der vorliegenden Analyse liegen in ihrer Verankerung in einem theoretischen

und methodischen Bezugsrahmen und einer umschriebenen Anfangs-Fragestellung, einer größeren Zahl an Interviews, die ein breiteres Spektrum an befragten Betroffenen in Bezug auf Hintergrundvariablen abdeckt, in dem Einsatz eines mehrköpfigen Forscherteams mit der Möglichkeit der Integration verschiedener Perspektiven, in ausführlicheren Patientenzitaten und in dem rhetorischen Rückgriff auf Extremtypen, wodurch eine prägnante Darstellung der Inhalte ermöglicht wurde. Die Publikation konnte den Prozess der Erkenntnisgewinnung nicht vollständig deutlich machen. Hier zeigt sich ein wesentlicher Nachteil der Darstellungsmöglichkeiten qualitativer Forschung in in ihrem Umfang stark begrenzten Veröffentlichungsmedien wie Journalartikeln. Die Darstellung des Forschungsprozesses, der eigenen Reflexion und Entwicklung von Forschungsfrage, Methode, Datenerhebung und Interpretation ist so platzintensiv, dass häufig nur das Endprodukt dargestellt wird. Dieses Merkmal teilen qualitative Forschungsberichte mit Artikeln aus der quantitativen Forschung, die entsprechende Prozessaspekte und -veränderungen im Verlauf einer Erhebung eher als Makel einer Untersuchung ansehen, die dennoch wichtig zu dokumentieren sind. Vor diesem Hintergrund wird nicht klar, inwieweit die aufgeführten Konstruktionen nicht vielmehr das Resultat der Vorannahmen der Forscher darstellten als das Ergebnis eines iterativen Analyseprozesses.

14.5 Wie lauten die Studienergebnisse?

Die Autoren stellen im Ergebnisteil zwei extreme Fälle dar, anhand derer sie zwei diskursive Konstruktionen mit Entsprechungen in der biomedizinischen Diskursen einführen (Patienten als Opfer der äußeren Umstände; Patienten, die für ihre Situation selber Verantwortung tragen). Sie betonen allerdings,

dass die meisten Patienten auf beide Formen der Diskurse in der Beschreibung von Ursachen und vom Umgang mit der Erkrankung zurückgreifen.

14.5.1 „Down to me“

Andy, ein 40-jähriger Mann unter medikamentöser Behandlung, trägt aus seiner Sicht die volle Verantwortung für die Entstehung des Diabetes. Er verweist ausschließlich auf Risikofaktoren, die unter seiner eigenen Kontrolle liegen: „... it was just gluttony [Völlerei] and no exercise whatsoever.“ (S. 100). Dadurch übernimmt er auch die volle Verantwortung für den Umgang mit der Erkrankung: „My support is just all from within.“ (S. 100). Er rekonstruiert die Erkrankung sogar als eine positive Erfahrung: „For the first time in my life I thought well I'm not big anymore. I really feel really good ... I think ideally was a good thing not a bad thing.“ (S. 100). Erfolge der Behandlung, u.a. drastische Gewichtsabnahme, Verschwinden des übermäßigen Durstgefühls und häufigen Urinierens, schreibt er seinen Handlungen zu, sie dienen dabei der Selbstwerterhöhung. Professionelle Unterstützung würde aus seiner Sicht erst dann wieder notwendig, wenn seine eigene Kontrolle über die Erkrankung versagen würde. Dass seine Ärztin auf seine Expertise im Umgang mit seiner Erkrankung baut und vertraut, bestärkt ihn in seiner Wahrnehmung. Der wahrgenommene Erfolg seines Selbstmanagements lässt ihn darüber spekulieren, inwieweit er noch tatsächlich diabetisch sei. Die ursprüngliche Diagnose stellt er jedoch nicht infrage.

14.5.2 „Up to them“

Sandra, ebenfalls 40 Jahre alt, sollte ihren Diabetes durch Ernährungsumstellung in den Griff bekommen. Sie stellt das andere

Extrem in Bezug auf die Erklärung und Umgang mit der Erkrankung dar: „It's in your system anyway you know, with what I've been told anyway, being hereditary and stuff like that.“ (S. 101). Die Ursachen ihrer Erkrankung sieht sie außerhalb ihrer Kontrolle (Vererbung, integraler Bestandteil ihres Körpersystems). Zudem mobilisiert sie die Meinungen von anderen, womit sie sich von der eigenen Verantwortlichkeit bzw. Schuldzuschreibungen freizusprechen sucht. Im Umgang mit der Erkrankung versteht sie Änderungen des Lebensstils und Medikamenteneinnahme nicht als zwei komplementäre, sich ergänzende Strategien, sondern vielmehr als nebeneinander stehende Alternativen: „I'm thinking well, as [my husband] says ‚Well you don't really want to be on tablets so you'll need to watch yourself and keep it [blood glucose] down'. I says ‚But if the tablets are going to keep me alive for the rest of my life I'd rather be on them'.“ (S. 101). Medikamente stellen vor diesem Hintergrund für Sandra die bevorzugte Option dar, ihre Erkrankung zu kontrollieren. Zwar weiß sie, welche Nahrungsmittel sie vermeiden sollte, aber ihre Entschlossenheit zur Ernährungsumstellung hat sich zum zweiten Messzeitpunkt weiter verringert. Professionelle Unterstützung hat für sie einen großen Stellenwert, obwohl sie die Ratschläge nur bedingt befolgt. Zwar sieht sie ihrem nächsten Kontakt mit der Krankenschwester mit Unbehagen entgegen, gleichzeitig hat dieser Kontakt auch eine entlastende Funktion für sie: „I came out feeling happier because I got the result I wanted you know. I'd been in and I'd let, I'd got it off my chest. I told her everything. That I'd been bad.“ Missbilligungen vonseiten der zuständigen Krankenschwester wurden in diesem Kontext sogar positiv gedeutet. Zwölf Monate nach der Erstbefragung ist sie noch weiter davon entfernt, Ernährungsratschläge einzuhalten. „I just thought ‚Och to hell' I'm going to enjoy myself ... And then I went to,

I went to the nurse and she gave me a row. So that was good.“ (S. 102). Die – unerfreuliche – Aufgabe des Umgangs mit der Erkrankung hat sie von sich geschoben. Sie würde sogar den Übergriff professioneller Mitarbeiter in ihre Privatsphäre in Kauf nehmen: „... [have health professionals] come in and just fill my cupboards and the fridge and the freezer with what I was to have.“ (S. 101).

14.5.3 Zusammengesetzte Positionen

Die Mehrzahl der Patientenaussagen lassen sich nicht einer der beiden o.g. Extrempositionen zuordnen. Die verursachenden Faktoren werden sowohl als kontrollierbar als auch als außerhalb der eigenen Kontrolle wahrgenommen. Die Vorstellung der eigenen Kontrollierbarkeit des Krankheitsverlaufs wird durch Erfahrungen konterkariert, dass trotz bester Bemühungen keine Besserung der Krankheitsparameter zu verzeichnen ist. Hierdurch erhält die Erkrankung in der Wahrnehmung der Betroffenen einen unkontrollierbaren Charakter, was zu bedeutenden Verunsicherungen auf Seiten der Befragten führt: „... it just seems to be that when I have a glass of Coca Cola it [blood glucose level] seems normal. So am I allowed a can every night?“ (S. 103). Patienten, deren Situation sich trotz Änderung des Lebensstils nicht verbesserte, konnten keine Bestätigung von professioneller Seite erwarten. Medikation wurde in diesem Zusammenhang zur Versinnbildlichung persönlichen Scheiterns, ein erfolgreiches Selbst-Management der Erkrankung durch eine Veränderung des Lebensstils zum unerreichbaren Ziel.

Die Autoren resümieren, dass trotz gleicher Informationen und Ratschläge, die den Patienten gegeben wurde, sich diese in ihren Interpretationen über die Ursachen der Erkrankung deutlich unterschieden. Die Vorstellungen der Patienten standen in engem Zusammenhang mit ihrem Umgang mit der Erkrankung und damit auch mit erfahrener

professioneller Unterstützung. Um angemessene Änderungen des Lebensstils für die Betroffenen zu ermöglichen, erscheint es notwendig, diese Vorstellungen von den Erkrankten zu erfahren und im Umgang mit den Betroffenen zu berücksichtigen.

14.6 Sind die Ergebnisse für die Behandlung meiner Patienten nützlich?

14.6.1 Ist mein eigenes klinisches Umfeld vergleichbar mit dem der Untersuchung?

Die Studie von Parry et al. [15] wurde im primärärztlichen Setting in Schottland durchgeführt. Patienten wurden in 16 Praxen bzw. in drei Krankenhäusern rekrutiert, was darauf verweist, dass Krankenhäuser dort ebenso einen Beitrag zur primärärztlichen Versorgung leisten. Einige Zitate im Ergebnisteil verweisen zudem darauf, dass dem Pflegepersonal in der Versorgung der Patienten eine bedeutende Rolle zukommt: Patientin Sandra beschrieb einen Termin mit einer „practical nurse“, bei dem sie Tabletten von ihr erhalten hat. Tatsächlich haben qualifizierte Pflegekräfte in Schottland eine zentrale Gate-keeper-Funktion in der primärmedizinischen Versorgung, ebenso werden Patienten weitervermittelt an Pflegekräfte mit entsprechenden Fachspezialisierungen. Alle Patienten haben eine vergleichbare Diabetes-Schulung erhalten – ein Ziel, das Sie für Ihre Patienten zwar anstreben, das aber doch in der Vergangenheit immer wieder an fehlenden Möglichkeiten und Kapazitäten gescheitert ist.

Es ist zu diskutieren, inwieweit diese Unterschiede in der Versorgung relevant für Ihr klinisches Problem sind. Sie fragen sich allerdings, inwieweit nicht eine Verteilung der Verantwortung auf mehrere Professionelle (Arzt, Pflege) dazu beitragen kann, dass die Patienten weniger einheitliche Erklärungs-

ansätze ihrer Erkrankung von professioneller Seite vermittelt bekommen und sich aus divergierenden Informationen eher das für sie Passendste herausgreifen. Diese Überlegung verdeutlicht Ihnen gleichzeitig, dass Sie selbst nicht wissen, aus welchen weiteren Quellen Ihre Patienten ihr Wissen über die Diabetes-Erkrankung schöpfen.

14.6.2 Können die Ergebnisse auf meinen Patienten angewendet werden?

Das primäre Ergebnis der vorliegenden Arbeit stellt die Art und Weise dar, wie Patienten mit Diabetes mellitus Typ 2 auf biomedizinische Diskurse für die Erklärung und den Umgang mit der Erkrankung rekurren. Auf dieser Ebene würden Sie keine grundlegenden Unterschiede zwischen der schottischen und der deutschen Kultur erwarten, anders als beispielsweise bei Studien aus dem asiatischen Kulturraum oder aus stark religiös geprägten Gesellschaften. Im Gegenteil: Sie haben eher das Gefühl, als sei in der Studie von Parry et al. etwas formuliert worden, das Ihnen selber schon bekannt war – Sie vermissen fast die „neue Erkenntnis“. Dieser Umstand spricht für die Übertragbarkeit der Ergebnisse auch auf Ihren Patienten. Zu berücksichtigen ist die Frage der Stabilität der Krankheitsmodelle im Krankheitsverlauf, da Ihr eigener Patient, Herr M., nicht mehr zu den neu erkrankten Personen zu zählen ist. Aus der Studie von Parry et al. [15] ist abzuleiten, dass die Erfahrungen der Patienten mit der eigenen Erkrankung einen eigenen Beitrag zur Entwicklung von Krankheitsvorstellungen leisten.

14.6.3 Wurden alle relevanten Dimensionen des klinischen Problems behandelt?

Der vorliegende Artikel fokussiert auf zwei Compliance-relevante Dimensionen des Patientenverhaltens: Vorstellungen über die Verursachung der Erkrankungen sowie Kontrollüberzeugungen. Eine systematische Übersichtsarbeit aus dem Jahr 2003 von Campbell et al. [10] fasst relevante Dimensionen aus verschiedenen Studien über Erfahrungen mit Diabetes und seiner Versorgung zusammen, in der das Konzept der Kontrolle einen zentralen Stellenwert einnimmt. Sie verweist auf die positive Bedeutung „strategischer Non-Compliance“, d.h. der bedachten und selektiven Anwendung medizinischer Ratschläge im Gegensatz zu einer „blinden Befolgung“. Dies ermöglichte den Patienten, ihren Diabetes zu kontrollieren, ein Gleichgewicht zwischen ihren Lebensqualitäts-Zielen und den Anforderungen der Erkrankung zu halten, ihren Blutzuckerspiegel zu verbessern und sich wohler zu fühlen. Als relevante Einflussfaktoren für diesen Zustand wurden identifiziert:

- ▲ Basiswissen über die Erkrankung,
- ▲ Anerkennung der Ernsthaftigkeit der Erkrankung vonseiten der Patienten,
- ▲ eine unterstützende Versorgung, die Informationen gibt, beim Selbst-Monitoring unterstützt sowie Verständnis – im Gegensatz zur Bewertung – für die eigenen Selbst-Management Strategien aufbringt,
- ▲ Zeit und Erfahrung in der Selbstbeobachtung eigener körperlicher Reaktionen,
- ▲ Vertrauen in die eigenen Handlungen und Beobachtungen auf Seiten der Patienten und
- ▲ eine weniger unterwürfige, sondern vielmehr hinterfragende Haltung gegenüber den professionellen Mitarbeitern.

In diesem Zusammenhang sollte auch darauf verwiesen werden, dass der Suchbegriff „Compliance“ zwar auf den ersten Blick passend scheint, er gleichzeitig entscheidende theoretische Entwicklungen in diesem wissenschaftlichen Feld vernachlässigt. Compliance wird als Ausdruck für ein paternalistisches Arztbild verstanden, wonach der Arzt den Standard setzt, nach dem der Patient sich richten muss, wenn er nicht als abweichend, unwillig oder defizitär gelten möchte. Um die gestärkte Rolle der Patienten in einem partnerschaftlichen Arzt-Patient-Verhältnis, in dem der Patient als primäre Entscheidungsinstanz und der Arzt als „expert adviser“ verstanden wird, zu berücksichtigen, werden sowohl die Begriffe „adherence“ als auch noch treffender „concordance“ verwendet [27]. Eine weiterführende Literaturrecherche müsste diese Begriffe in die Suche integrieren.

14.6.4 Verändern die Erkenntnisse den Umgang mit meinem Patienten?

Die Lektüre und Analyse der Arbeit von Parry et al. [15] hat Ihnen deutlich vor Augen geführt, dass der Umgang der Patienten mit ihrer Erkrankung in entscheidender Art und Weise von den Vorstellungen der Patienten über die Entstehung der Erkrankung und ihrer möglichen Beeinflussung geprägt ist. Diese Vorstellungen können trotz einheitlicher Informationsvermittlung von Ihrer Seite sehr stark variieren. Ihr Umgang mit Ihren Patienten wird sich vor diesem Hintergrund dahingehend verändern, dass Sie Ihren Patienten mehr Möglichkeiten einräumen wollen, über ihre Erkrankung und ihren Umgang damit zu sprechen, um daran anknüpfend weitere Strategien zu besprechen. Ebenso werden Sie mehr auf Formulierungen achten, die von Patientenseite aufzeigen, wen sie für den Umgang mit der Erkrankung verantwortlich halten.

14.7 Auflösung des klinischen Szenarios

Den nächsten anstehenden Kontrolltermin von Herrn M. lassen Sie nicht wie sonst in die unruhigen frühen Vormittagsstunden, sondern auf den Nachmittag, nach seinem Dienstschluss legen, damit beide Seiten nicht in unnötigen Zeitdruck kommen. Sie nehmen sich vor, herauszufinden, wie Herr M. seine Erkrankung versteht, und mehr darüber zu erfahren, wie er mit ihr im Alltag umgeht – ohne gleich in die Rolle zu verfallen, gut gemeinte Ratschläge geben zu müssen. Sie erfahren, dass Herrn M.s Vorstellungen über seine Erkrankung entscheidend durch seinen Onkel geprägt wurden, der in etwas späteren Jahren als Herr M. selbst an Diabetes erkrankte. Ihm sei klar geworden, dass seine Erkrankung „in der Familie liege“ und er wenig gegen ihre Entstehung machen könne. Sein Onkel habe zudem nicht viel an seinem Lebensstil verändert, er sei trotzdem über 70 Jahre alt geworden und „nicht an Diabetes, sondern am Herzinfarkt gestorben“. Herr M. habe zudem die Erfahrung gemacht, dass sich sein Blutzucker nicht verbessert habe, trotz seiner Bemühungen, seine Lebensgewohnheiten zu verändern. Im Gegenteil hätten Sie ihn gerade zu einer Zeit sehr gescholten, in der er sich viel Mühe gegeben hätte. So habe er sich damit arrangiert, dass er ein relativ normales Leben ohne weitere Beschwerden führen könne, dafür aber immer wieder Ermahnungen von Ihrer Seite in Kauf nehmen müsse. Spritzen lehnt Herr M. auch deswegen ab, da sein Onkel einige Jahre vor seinem Tod auch angefangen hatte zu spritzen, und er das Spritzen mit einer Verschlechterung seines Zustandes in Zusammenhang bringe.

Ein Ziel Ihrer Bemühungen haben Sie erreicht: einen besseren Zugang zu Ihrem Patienten. Von Ihrem weiteren Ziel, Herrn M. von der Notwendigkeit einer intensivierten Insulin-Therapie zu überzeugen, nehmen Sie vorerst Abstand. Stattdessen möchten Sie Herrn M. ermöglichen, Zusammenhänge zwischen seinem Ernährungs- und Bewegungsverhalten und dem

Blutzuckerspiegel erfahrbar zu machen und dadurch die eigenen Kontrollmöglichkeiten realistischer einschätzen zu lernen. Sie besprechen die Möglichkeit, vor und nach bestimmten Aktivitäten den Blutzucker im Alltag zu messen und dies in einem Diabetes-Tagebuch kurz festzuhalten. Sie vereinbaren, in den nächsten Wochen sich nicht zum Ziel zu setzen, den HbA1c-Wert zu senken, sondern gemeinsam etwas über die Schwankungen des Blutzuckers bei Herrn M. in Abhängigkeit von bestimmten Aktivitäten zu lernen. Herr M. wird sich bis zu Ihrem nächsten Termin um die Beschaffung eines Blutzuckermessgeräts kümmern. Sie sagen zu, sich über die Möglichkeiten strukturierter Blutzucker-Selbstwahrnehmungstrainings in Ihrer Umgebung zu informieren.

14.8 Ausblick

„Evidence does not make decisions, people do.“ [28]

Gerade Veröffentlichungen aus dem Kontext quantitativer Forschung neigen dazu, qualitativen Studien primär eine Hypothesen generierende Rolle zuzuweisen bzw. sie zur Unterstützung der Interpretation quantitativer Studienergebnisse heranzuziehen (z.B. [29]). Hier werden die Potenziale und die Breite der qualitativen Forschung allerdings deutlich unterschätzt, insbesondere im Kontext der Evidenzbasierten Medizin. Entscheidungsmodelle der EbM, z.B. von Haynes et al. [28] (vgl. Kap. 2) verstehen externe Evidenz als **eine** wichtige Säule der klinischen Entscheidungsfindung. Als weitere Säule haben Patientenpräferenzen und -handlungen einen integralen Bezug zur qualitativen Forschung. Eine der großen Stärken qualitativer Forschung besteht gerade darin, die Sichtweisen von Patienten transparent machen zu können und ihre Handlungen und Entscheidungen erklärbarer. Z.B. wurde mit „DIPeX“ (database of patients' experiences, www.dipeX.org; [30]) in England eine Initiative ent-

wickelt, um Patientenerfahrungen („patient accounts“) über ihre Erkrankungen, Behandlungen und Folgen für Patienten, Angehörige, aber auch Professionelle zugänglich zu machen. Ein weiteres wichtiges Feld qualitativer Studien bezieht sich auf die Frage, auf welche Weise der Kliniker seine Entscheidungen trifft, z.B. wie er externe Evidenz in seine klinischen Entscheidungen integriert. D.h. die Implementierung der EbM im Denken und Handeln der Ärzte [22], insbesondere das Aufdecken von möglichen Hinderungsgründen für das Nutzen externer Evidenz, stellt ein originäres Forschungsfeld qualitativer Studien dar. Nicht zuletzt können – insbesondere im Kontext chronischer Erkrankungen – Studien über Krankheitsverläufe wichtige Hinweise für eine ausführliche Anamnese geben, indem der Arzt auf für die Behandlung zentrale Aspekte hingewiesen wird, die im Alltag aus dem Blick geraten, z.B. die Rolle und Bedeutung der Angehörigen für die Begleitung, Versorgung und den Krankheitsverlauf des Betroffenen.

Literatur

- [1] Pope C, Mays N, Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *BMJ*. (1995), 311, 42–45
- [2] Green J, Britten N, Qualitative research and evidence based medicine. *BMJ*. (1998), 316, 1230–1232
- [3] Morse M et al. (2001) The nature of qualitative evidence. Sage, Thousand Oaks
- [4] Jones R, Why do qualitative research? *BMJ*. (1995), 311, 2
- [5] Malterud K, Qualitative research: standards, challenges, and guidelines. *Lancet*. (2001), 358, 483–488
- [6] Smith LK, Pope C, Botha JL, Patients' help-seeking experiences and delay in cancer presentation: a qualitative synthesis. *Lancet*. (2005), 366, 825–831
- [7] Giacomini MK, Cook DJ, Users' guides to the medical literature: XXIII. Qualitative research in health care A. Are the results of the study valid? *Evidence-Based Medi-*

- cine Working Group. *JAMA*. (2000), 284, 357–362
- [8] Giacomini MK, Cook DJ, Users' guides to the medical literature: XXIII. Qualitative research in health care B. What are the results and how do they help me care for my patients? Evidence-Based Medicine Working Group. *JAMA*. (2000), 284, 478–482
- [9] Bradley EH et al., A qualitative study of increasing beta-blocker use after myocardial infarction: Why do some hospitals succeed? *JAMA*. (2001), 285, 2604–2611
- [10] Campbell R et al., Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. *Soc Sci Med*. (2003), 56, 671–684
- [11] Harden A et al., Applying systematic review methods to studies of people's views: an example from public health research. *J Epidemiol Community Health*. (2004), 58, 794–800
- [12] Wong SS, Wilczynski NL, Haynes RB, Developing optimal search strategies for detecting clinically relevant qualitative studies in MEDLINE. *Medinfo*. (2004), 11, 311–316
- [13] Vinter-Repalust N, Petricek G, Katic M, Obstacles which patients with type 2 diabetes meet while adhering to the therapeutic regimen in everyday life: qualitative study. *Croat Med J*. (2004), 45, 630–636
- [14] Holmstrom IM, Rosenqvist U, Misunderstandings about illness and treatment among patients with type 2 diabetes. *J Adv Nurs*. (2005), 49, 146–154
- [15] Parry O et al., Issues of cause and control in patient accounts of Type 2 diabetes. *Health Educ Res*. (2006), 21, 97–107
- [16] Greenhalgh T, Taylor R, Papers that go beyond numbers (qualitative research). *BMJ*. (1997), 315, 740–743
- [17] Lambert H, McKeivitt C, Anthropology in health research: from qualitative methods to multidisciplinary. *BMJ*. (2002), 325, 210–213
- [18] Flick U et al. (2004) *Qualitative Forschung. Ein Handbuch*. Rowohlt, Reinbek bei Hamburg
- [19] Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P. Qualitative research methods in health technology assessment: a review of the literature. *Health Technology Assessment 1998 Vol. 2: No. 16. (Executive Summary)*. <http://www.hta.nhsweb.nhs.uk/execsumm/SUMM216.HTM> (24.11.06)
- [20] Rost J, Drei Thesen zum Konzept qualitativer Forschungsmethoden. *Z Didakt Naturwiss*. (1998), 4, 35–42
- [21] Barbour RS, Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? *BMJ*. (2001), 322, 1115–1117
- [22] Malterud K, The art and science of clinical knowledge: evidence beyond measures and numbers. *Lancet*. (2001), 358, 397–400
- [23] Sandelowski M, Barroso J, Reading qualitative studies. *Int J Qual Meth*. (2002), 1, Article 5
- [24] Popay J, Qualitative research and process evaluation in systematic reviews: the methodological research agenda. *The Cochrane Collaboration Methods Groups Newsletter*, Volume 4. <http://www.cochrane.org/newslett/mg4-1.htm#Qualitative%20research%20and%20process%20evaluation%20in> (26.11.06)
- [25] Steinke I (1999) *Kriterien qualitativer Forschung. Ansätze zur Bewertung qualitativ-empirischer Sozialforschung*. Juventa-Verlag, Weinheim
- [26] Mays N, Pope C, Qualitative research in health care. Assessing quality in qualitative research. *BMJ*. (2000), 320, 50–52
- [27] Vermeire E et al., The adherence of type 2 diabetes patients to their therapeutic regimens: a qualitative study from the patients' perspective. *Pract Diabetes Int*. (2003), 20, 209–214
- [28] Haynes RB, Devereaux PJ, Guyatt GH, Physicians' and patients' choices in evidence based practice. *BMJ*. (2002), 324, 1350
- [29] Bastian H et al., *Methoden des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)*. Version 1.0 vom März 2005. http://www.iqwig.de/download/2005_03_01_IQWiG_Methoden.pdf (24.11.06)
- [30] Herxheimer A et al., A database of patients' experiences (DIPex): Patienten berichten von ihren Krankheitserfahrungen. *Z Allg Med*. (2001), 77, 323–327

15 Kritische Bewertung von Leitlinien

Henning Thole, Frank Thalau, Günter Ollenschläger, Ina Kopp, Monika Lelgemann

Kernaussagen

- ▲ Leitlinien kondensieren und werten medizinisches Wissen (Evidenz) zu speziellen Fragestellungen und geben Empfehlungen für das ärztliche Handeln.
- ▲ Die Zuverlässigkeit von Leitlinien kann variieren und bedingt die Notwendigkeit einer kritischen Bewertung ihrer Qualität.
- ▲ Strukturierte Bewertungsinstrumente (Checklisten) sind ein wichtiges Hilfsmittel, denn sie fassen formale und methodische Kriterien zur Leitlinienbewertung in übersichtlicher Weise zusammen.
- ▲ Mit dem Deutschen Leitlinien-Bewertungs-Instrument (DELBI) [1, www.delbi.de] steht ein für den deutschen Versorgungsbereich angepasstes Bewertungsinstrument zur Verfügung.

15.1 Einleitung

In der Evidenzbasierten Medizin dient die kritische Bewertung von Informationen im Wesentlichen der Überprüfung im Hinblick auf ihre Verlässlichkeit, Relevanz und Anwendbarkeit (s. hierzu auch Kap. 2).

Dieses Vorgehen gilt prinzipiell auch für die Leitlinienbewertung. Ziel ist es, systematisch anhand expliziter und vorab definierter Kriterien vorzugehen, um sich einen Eindruck von der methodischen Qualität und der Belastbarkeit der Informationen bilden zu können. Dabei sollte man sich bewusst sein, dass jede Bewertung subjektive Elemente enthält. Die Entscheidung, ob man am Ende einer Leitlinie vertrauen soll oder

nicht, ist jedoch besser begründet, wenn – anstelle einer impliziten Bewertung – ein systematisches und transparentes Vorgehen anhand nachvollziehbarer Kriterien gewählt wird. Dabei kann eine unabhängige Bewertung durch mehrere Personen und eine strukturierte Diskussion entlang der Kriterien zu besser begründeten Entscheidungen beitragen.

15.2 Das Szenario

Sie sind als Hausarzt niedergelassen und seit kurzer Zeit Moderator eines ärztlichen Qualitätszirkels in Ihrer Gegend. Aufgrund von Neuerungen in der Therapie der chronischen koronaren Herzkrankheit, die sich seit Monaten durch die Fachpresse ziehen, haben Sie nach deutschen Leitlinien zu dem Thema gesucht und möchten auf einem der nächsten Treffen Ihres Qualitätszirkels einen Überblick über die leitliniengerechte Therapie geben.

Dazu haben Sie im Internet nach Leitlinien recherchiert und verschiedene Leitlinien unterschiedlichen Umfanges gefunden, die in den letzten vier Jahren erschienen sind. Sie fragen sich, welche Leitlinie Sie nun verwenden sollen, woran Sie die Zuverlässigkeit erkennen und wie Sie Ihre Auswahl ggf. begründen können. Bei Ihrer Leitliniensuche haben Sie auf der Seite www.leitlinien.de bereits ein deutschsprachiges Bewertungsinstrument für Leitlinien gefunden und Sie beschließen, sich dieses näher anzusehen.

15.3 Qualitätskriterien von Leitlinien

Für zukünftige Anwender bemisst sich die Qualität einer Leitlinie nicht nur daran, wie gut sie das aktuelle Wissen zusammenstellt, sondern auch in der Klarheit und Umsetzbarkeit der Empfehlungen. Leitlinien sind also mehr als Evidenzberichte oder Quellen „aufbereiteter Evidenz“, denn sie geben konkrete Handlungsempfehlungen.

Die Vielzahl internationaler und nationaler Leitlinien macht eine kritische Auswahl erforderlich, zumal oft mehrere Leitlinien zu einem Krankheitsbild vorliegen, die unterschiedliche Empfehlungen zu gleichen Fragestellungen enthalten können.

Es liegen empirische Studien vor, die zeigen, dass die Erfüllung methodisch-inhaltlicher Forderungen in Leitlinien deren Anwendung und damit die „Wirksamkeit“ der Leitlinien bestimmt [2–4]. Bereits Anfang der 1990er-Jahre wurden Qualitätskriterien für Leitlinien beschrieben [5–7], die in Tabelle 15.1 dargestellt sind.

15.4 Bewertung formaler und methodischer Aspekte von Leitlinien

Für die formale und methodische Bewertung von Leitlinien, d.h. die Bewertung des Erstellungsprozesses und der Art der Darstellung der Empfehlungen, stehen verschiedene Instrumente zur Verfügung. International etabliert ist das validierte AGREE-Instrument (Appraisal of Guidelines Research and Evaluation in Europe) [8]. Hierin werden die oben genannten Qualitätskriterien in sechs Gruppen (so genannten Domänen) zusammenfasst und für jede Domäne eine bestimmte Anzahl so genannter Statements („Fragen“) formuliert, anhand derer die Qualität bewertet werden kann (s. Tab. 15.6).

Für den deutschsprachigen Raum existiert das – vom AGREE-Instrument adaptierte

und weiterentwickelte – Deutsche Instrument zur methodischen Leitlinien-Bewertung (DELBI), welches u.a. um eine zusätzliche siebte Domäne ergänzt wurde, die Aspekte der Übertragbarkeit auf das deutsche Gesundheitssystem umfasst. Tabelle 15.2 stellt die Domänen in Kurzfassung zusammen, Tabelle 15.6 enthält die insgesamt 29 dazugehörenden Statements.

Die wesentliche Weiterentwicklung von DELBI gegenüber dem ursprünglichen AGREE-Instrument liegt in der Ausformulierung von Beurteilungskriterien, mit deren Hilfe auf einer Skala von 1–4 der Grad der Erfüllung jedes Qualitätskriteriums geprüft werden kann. Im Gegensatz zur reinen Ja/Nein-Abfrage anderer Checklisten wird somit eine Aussage getroffen, die nicht nur über das Vorhandensein eines Qualitätsmerkmals, sondern auch über dessen Ausprägung Auskunft gibt.

15.5 Bewertung inhaltlicher Aspekte von Leitlinien

Derzeit existiert national und international kein empirisch abgesichertes oder konsentiertes Verfahren, mit dem die inhaltliche „Korrektheit“ von Leitlinienempfehlungen überprüft werden könnte. Zwischen der formalen und methodischen Qualität von Leitlinien einerseits und ihrer inhaltlichen Qualität andererseits besteht kein garantierter Zusammenhang [9].

Dennoch ist eine näherungsweise Bewertung der Qualität von Leitlinien und den Empfehlungen möglich – insbesondere dann, wenn das Vorgehen bei der Formulierung von Leitlinien-Empfehlungen entsprechend den Prinzipien der Evidenzbasierten Medizin transparent dargestellt wurde.

Die Schwierigkeit einer Überprüfung der inhaltlichen Angemessenheit und der Anwendbarkeit der Empfehlungen einer Leitlinie liegt wesentlich darin begründet, dass bei

Tab. 15.1 Qualitätskriterien von Leitlinien

• Transparenz	• Klarheit, Eindeutigkeit
• Gültigkeit	• Dokumentation der Leitlinienentwicklung
• Zuverlässigkeit und Reproduzierbarkeit	• Planmäßige Überprüfung
• Multidisziplinäre Entwicklung	• Überprüfung der Anwendung
• Anwendbarkeit	• Kosten-Nutzen-Verhältnis
• Flexibilität	• Verfügbarkeit der Leitlinie

Tab. 15.2: Gemeinsame AGREE- und DELBI-Domänen und zusätzliche DELBI-Domäne

1. Geltungsbereich und Zweck	
2. Beteiligung von Interessengruppen	
3. Methodologische Exaktheit der Leitlinienentwicklung	
4. Klarheit der Präsentation	
5. Generelle Anwendbarkeit	
6. Redaktionelle Unabhängigkeit	
7. Anwendbarkeit im deutschen Gesundheitswesen	Zusätzlich in DELBI

Tab. 15.3: Beispiel aus DELBI (Domäne 2, Statement 4)

Zusammensetzung der Leitliniengruppen
4. Die Entwicklergruppe der Leitlinie schließt Mitglieder aller relevanten Berufsgruppen ein.
Die Bereitschaft, Leitlinien anzuwenden, steigt in dem Maße, in dem der Erstellungsprozess transparent und die Entwicklergruppe repräsentativ in Bezug auf die Anwendergruppe sind. In die Entwicklung einer Leitlinie sollten daher alle von der Thematik betroffenen Fachgruppen, aber auch weitere Interessengruppen, einbezogen worden sein. Eine Einbeziehung kann direkt als Autor oder Teilnehmer am Konsensusverfahren, durch ein Review-Verfahren oder eine öffentliche Diskussion vor Inkrafttreten der Leitlinie geleistet werden. Die betroffenen Fachgruppen sind in der Regel diejenigen, die im Kontext der Leitlinie adressiert werden, darüber hinaus aber ggf. auch Gruppen, die an Schnittstellen benötigt werden (z.B. Rettungskräfte/Notärzte bei der Herzinfarktversorgung, Pflegepersonal, medizinische Hilfsberufe).
Das Statement muss mit „Trifft überhaupt nicht zu“ („1“) beantwortet werden, wenn keine Angaben über die Zusammensetzung der Entwicklergruppe gemacht werden oder wenn nur Autoren mit Namen genannt werden und so eine Zuordnung zu Fachgruppen oder Funktionsbereichen nicht möglich ist.
Für eine Beantwortung mit „2“ ist ergänzend zu den Namen die eindeutige Benennung der Fach- oder Interessensgruppen, die die jeweiligen Autoren repräsentieren, notwendig.
Für eine Beantwortung mit „3“ wird darüber hinaus aus der Benennung klar, dass alle im Kontext der Leitlinie adressierten Fachgruppen an der Entwicklung der Leitlinie beteiligt waren.
Für eine Antwort mit „Trifft uneingeschränkt zu“ („4“) ist ergänzend die Angabe über Art und Weise der Einbeziehung von nicht direkt im Kontext der Leitlinie adressierten Fach- und Interessensgruppen notwendig (z.B. Pflegeberufe bei einer rein ärztlichen Leitlinie). Diese Darlegung kann in der Leitlinie selbst oder in einem Leitlinien-Report (siehe Kriterium 29) erfolgen, wobei die Existenz eines Leitlinien-Reports klar aus der Leitlinie hervorgehen muss.

Tab. 15.4: Beispiel aus DELBI (Domäne 6, Statement 23)

Angaben zu Interessenkonflikte
23. Interessenkonflikte von Mitgliedern der Leitlinienentwicklungsgruppe wurden dokumentiert.
Jedes Mitglied einer Leitlinienentwicklungsgruppe hat verschiedenste berufliche und private Interessen und Verbindungen und kann deswegen in Interessenkonflikte kommen (z.B. wenn in einer Leitlinie von eigenen Forschungsergebnissen abweichende Positionen empfohlen werden). Dies trifft beispielsweise auch zu, wenn ein Autor auf dem von der Leitlinie betroffenen Gebiet wissenschaftlich arbeitet und dabei von einer pharmazeutischen Firma finanziell unterstützt wird.
Das Statement muss mit „Trifft überhaupt nicht zu“ („1“) beantwortet werden, wenn in der Leitlinie keine Angaben zur Interessenkonflikt-Darlegung enthalten sind.
Für eine Beantwortung mit „2“ ist es erforderlich, dass die Leitlinie mit einer globalen Erklärung die fehlende Beeinflussung durch die finanzielle Unterstützung darlegt.
Für eine Beantwortung mit „3“ ist es erforderlich, dass in der Leitlinie eine Darlegung zur Abfrage einer möglichen Beeinflussung mit Angabe der abgefragten Aspekte (z.B. in Form eines Formblattes) enthalten ist.
Für eine Antwort mit „Trifft uneingeschränkt zu“ („4“) ist es erforderlich, dass das Verfahren zur Identifizierung möglicher Interessenkonflikte einschließlich der Ergebnisse des Verfahrens explizit dargelegt ist. Dies kann in der Leitlinie selbst oder in einem Leitlinien-Report (s. Kriterium 29) erfolgen, wobei die Existenz eines Leitlinien-Reports klar aus der Leitlinie hervorgehen muss. Die Beschreibung im Leitlinien-Report muss sich in Bezug auf die hier geforderte Darlegung individuell auf die vorliegende Leitlinie beziehen, eine prinzipielle Darlegung ist nicht ausreichend.

der Formulierung der Empfehlungen einer Leitlinie Werturteile eine erhebliche Rolle spielen. Während die Bewertung einzelner Studien in der Regel eher unstrittig und nachvollziehbar ist, ist dies bei den Empfehlungen und der Zuordnung von Empfehlungsgraden zu den einzelnen Empfehlungen (so genannte Graduierung) deutlich schwieriger, da hier eine entscheidende Wertung durch die Autoren u.a. in Bezug auf die Ziele und den Adressaten-/Anwendungsbereich der Leitlinie erfolgt. Bei der Ableitung von Empfehlungen aus der Literatur handelt es sich um ein mehrstufiges Verfahren.

Bei dieser Ableitung sollte auf allen Stufen nachvollziehbar sein,

- ▲ ob eine systematische und vollständige Recherche die jeweils „best verfügbare Evidenz“ (Informationsquellen, v.a. Studien) gefunden hat.
- ▲ ob die Bewertung der jeweiligen Evidenz anhand nachvollziehbarer methodischer Kriterien erfolgte.

- ▲ ob die inhaltliche Ableitung der Empfehlung aus der zuvor identifizierten und ausgewählten Evidenz korrekt ist. Für den Fall, dass Diskrepanzen zwischen Evidenz und Empfehlung vorliegen, sollten die Gründe hierfür für den Nutzer nachvollziehbar dargelegt sein.
- ▲ ob die Empfehlungsgrade einen nachvollziehbaren Zusammenhang mit der Evidenzeinstufung der jeweils zitierten Literatur haben, d.h. die Gründe für eine Herauf- bzw. Herabstufung der Empfehlungsgrade („Evidenz-Shifting“) in Relation zur zitierten Literatur dargelegt sind.

Um die methodische Qualität einer Leitlinie beurteilen zu können, ist es daher entscheidend, dass alle Wertungen, die bei der Erarbeitung einer Leitlinie vorgenommen wurden, möglichst transparent dargelegt worden sind. Die „inhaltliche Korrektheit“ einer Leitlinie lässt sich dann – allerdings nur indirekt – daran bemessen, ob die Übertragung

der Evidenz in die Empfehlungen nachvollziehbar und – auch in Bezug auf den jeweiligen Anwendungs-Zusammenhang – plausibel ist.

Leitlinien, die auf einem evidenzbasierten Erstellungsprozess aufbauen, haben eine größere Wahrscheinlichkeit, inhaltlich richtige Empfehlungen zu enthalten, weil die best verfügbare Evidenz systematisch gesucht, aufbereitet und ausgewählt wurde.

Mit der Frage nach der „inhaltlichen Korrektheit“ wird das Problem „angesprochen, dass – auch in Leitlinien, deren Methodik vielen der geforderten Qualitätskriterien entspricht – Empfehlungen enthalten sein können, die nicht auf die Patienten-Zielgruppe passen, die in der Leitlinie adressiert ist. Dies kann zum Beispiel daran festgestellt werden, dass

- ▲ es Abweichungen zwischen Evidenz und Empfehlung in Bezug auf Alter oder Komorbidität oder Prävalenz gibt [10; 11], oder
- ▲ die Ableitung der Evidenz aus der Literatur fehlerhaft ist [12].

Dabei ist die „inhaltliche Korrektheit“ von Leitlinienempfehlungen nicht mit ihrer Relevanz und Anwendbarkeit gleichzusetzen. Selbst eine inhaltlich richtige und relevante Empfehlung kann dennoch nicht anwendbar sein, wenn etwa bestimmte finanzielle Mittel zur Umsetzung der Empfehlung fehlen. Solche Einschränkungen der Umsetzbarkeit sollten von den Leitlinienautoren aufgeführt werden, da sie ein großes Implementierungshindernis darstellen.

Weitere Möglichkeiten zu Überprüfung der inhaltlichen Qualität von Leitlinien bieten Begutachtungen durch nicht an der Leitlinien-Erstellung beteiligte Fachexperten (Peer-Review-Verfahren) oder kontinuierliche Kommentierungsmöglichkeiten der Leitlinie durch die Öffentlichkeit, deren Ergebnisse in die (Weiter-)Entwicklung der Leitlinie bzw. deren Aktualisierung einfließen

(„Living Guideline“). Die Veröffentlichung aller Kommentare, z.B. über eine Internetseite, kann dabei als ein Verfahren mit besonders hoher Transparenz angesehen werden.

Letztlich ist die inhaltliche Qualität einer Leitlinie nur über die Beantwortung der Frage überprüfbar, ob sich die Qualität der medizinischen Versorgung und der erzielten Versorgungsergebnisse durch ihre Einhaltung tatsächlich verbessert. Zur Beurteilung des Nutzens einer Leitlinie wird man also ihre methodische Qualität, die Richtigkeit der Einführungsentscheidung, die Verbreitung, die Umsetzung und ihre Auswirkungen auf Strukturen, Prozesse und Ergebnisse der Versorgung gemeinsam bewerten müssen.

15.6 Anwendung von DELBI

Je nach Bedarf und verfügbaren zeitlichen Ressourcen für die Leitlinien-Bewertung können die Statements von DELBI in unterschiedlicher Weise abgearbeitet werden:

- ▲ Die gezielte Bearbeitung nur einzelner Statements als Filter- oder „Knock-Out-Kriterien“ kann gewählt werden, wenn bestimmte formale oder methodische Aspekte eine Grundbedingung für die weitere Berücksichtigung einer Leitlinie darstellen. So könnten z.B. die Multidisziplinarität der Leitlinien-Autorengruppe oder Aspekte der Evidenzbasierung (z.B. die Darlegung der systematischen Literaturrecherche zu einzelnen Fragenstellungen) solche Ausschluss- bzw. Einschlusskriterien darstellen.
- ▲ Um die Übertragbarkeit ausländischer Leitlinien zu prüfen, kann die Domäne 7 (Anwendbarkeit im deutschen Gesundheitssystem) im Sinne eines Screening-Instruments eingesetzt werden, ehe eine weitergehende Bewertung erfolgt.
- ▲ Die vollständige Bewertung aller 29 Statements in den sieben Domänen sollte vor

der Anwendung/Implementierung einer Leitlinie in einer bestimmten Institution (z.B. Klinik, Abteilung, Ärztenetz) erfolgen.

Aus der Sicht von Leitlinienautoren kann DELBI eine Orientierung für die Leitlinienerstellung geben, da die in der Checkliste abgefragten Aspekte das „Pflichtenheft“ von Leitlinienautoren darstellen, wobei die Umsetzung einzelner methodischer Schritte in übersichtlicher Weise in einem Methodenreport zusammengefasst werden sollte.

Für die Auswahl einer Leitlinie können die Leitlinienbewerter anhand der Antwortstufen 1–4 vorab festlegen, welche Mindestanforderungen eine Leitlinie erfüllen sollte, wenn sie zur Anwendung in ihrem Versorgungsbereich infrage kommen soll. Die Fokussierung auf einzelne Aspekte der Qualität von Leitlinien durch die Berücksichtigung nur einzelner Statements oder Domänen sollte nur mit dem Ziel einer groben Sortierung angewandt werden.

Die vollständige Beurteilung der formalen und methodischen Qualität einer Leitlinie erfordert die Beantwortung aller Statements aus allen Domänen durch – optimalerweise mehrere, gemäß DELBI mindestens zwei – Bewerter.

Für die praktische Bewertungsarbeit ist es wichtig, sich das gesamte zur Leitlinie gehörende Material zu beschaffen, da ein Großteil der Dokumentation zu den von den Autoren getroffenen Entscheidungen häufig nur in begleitenden Leitlinien- oder Methodenreports dargelegt ist.

Die Bewertung von Leitlinien mit DELBI ergibt im Ergebnis ein Bild, mit dem das formale und methodische Profil einer Leitlinie erfasst werden kann. Einschränkend ist hinzuzufügen, dass nur beurteilt werden kann, was seitens der Leitlinienautoren bzw. seitens der Herausgeber dargelegt wird. Insofern prüft DELBI streng genommen lediglich die so genannte Berichtsqualität der bei der Erstellung angewandten Arbeitsmethodik.

15.7 Clearingverfahren

Unter Clearingverfahren allgemein werden „aufgearbeitete Zusammenstellungen verschiedener Quellen“ [13] verstanden, die für ein bestimmtes Fachgebiet oder für mehrere Gebiete erarbeitet werden.

Auf dem Gebiet der Leitlinien wurden in den letzten Jahren zur Durchführung von Clearingverfahren so genannte Clearing-Stellen („Guideline Clearinghouses“) in verschiedenen Ländern eingerichtet. Ihr Ziel ist es, dem Nutzer den Zugang zu Leitlinien zu erleichtern, Transparenz bezüglich der Qualität herzustellen und eine Orientierung hinsichtlich der Leitlinienqualität anzubieten.

Bei den Clearing-Stellen kommen unterschiedliche Verfahren zur Anwendung, z.B.:

▲ Leitlinien-Register ohne Darlegung bestimmter Qualitätskriterien:

zum Beispiel die International Guideline Library von G-I-N (<http://www.g-i-n.net>)

▲ Leitlinien-Register selbstbewerteter Leitlinien:

die Qualitätsbewertung wird durch die Herausgeber/Autoren selbst durchgeführt (Beispiele: AWMF-Leitlinien-Datenbank (<http://www.awmf-online.de>) und US Guideline Clearinghouse (<http://www.guideline.gov>))

▲ Leitlinien-Register fremdbewerteter Leitlinien:

die Bewertung erfolgt in Form eines standardisierten Peer-Review-Verfahrens durch externe Experten, zum Beispiel: Deutsches Leitlinien-Clearingverfahren beim ÄZQ (<http://www.leitlinien.de/clearingverfahren>).

Im Folgenden soll kurz auf das Deutsche Leitlinien-Clearingverfahren eingegangen werden, dessen Ziel es war, die Anforderungen an eine gute Leitlinie zu dem behandelten Thema für zukünftige Leitlinienautoren darzulegen [14].

Zwischen 1999 und 2005 wurden zu 15 Krankheitsbildern insgesamt 230 Leitlinien

unterschiedlicher nationaler und internationaler Organisationen formal und methodisch mittels einer Checkliste bewertet [15]. Diese Bewertung wurde unabhängig von jeweils zwei Personen vorgenommen, im Fall von Dissens wurde eine dritte Person konsultiert.

Darüber hinaus wurde mithilfe multidisziplinär zusammengesetzter Expertengruppen, so genannter Fokus-Gruppen, die inhaltliche Angemessenheit der in den Leitlinien enthaltenen Empfehlungen diskutiert. An den Clearing-Verfahren zu den beiden onkologischen Themen „Mammakarzinom“ und „Kolorektales Karzinom“ wirkten zusätzlich Patientenvertreter mit.

Das Deutsche Leitlinien-Clearingverfahren stellt für Deutschland den bislang umfassendsten Ansatz dar, die methodische Qualität und die inhaltliche Angemessenheit von Leitlinien für das deutsche Gesundheitswesen zu bewerten.

15.8 Auflösung des Szenarios

Sie berichten Ihren Kolleginnen und Kollegen über die vier von Ihnen gefundenen Leitlinien und über die Möglichkeiten, eine geeignete, methodisch möglichst gute Leitlinie zu identifizieren. Hierzu stellen Sie die Beurteilungskriterien von DELBI vor. Sie schlagen vor, die Leitlinien gemeinsam mit allen Teilnehmern des Qualitätszirkels zu bewerten. Abweichende Bewertungsergebnisse diskutieren Sie gemeinsam.

Zwei der Leitlinien fallen durch sehr hohe Punktzahlen in der Domäne 2 (Beteiligung von Interessengruppen) und der Domäne 3 (Methodologische Exaktheit der Leitlinienentwicklung) auf.

Sie entscheiden sich, die weitere Auswahl auf diese beiden Leitlinien einzuzugrenzen. Eine von diesen beiden Leitlinien verfügt über umfangreiches Begleitmaterial für die Anwendung, einen Methodenreport und Patienteninformationsmaterial. Da die Einschätzung der Leitlinienqualität für beide annähernd gleich ist, beschließen

Sie, die Leitlinie mit dem umfangreicheren Zusatzmaterial für die weitere inhaltliche Diskussion heranzuziehen, da sie die Arbeit in der Praxis Ihrer Einschätzung nach besser unterstützt.

Da die Gesamtbewertung der vier gefundenen Leitlinien insgesamt recht lange gedauert hat, beschließen Sie, für künftige Entscheidungs-/Auswahlverfahren die Domäne 3, die Ihnen einen Überblick über die Erstellungsqualität der Leitlinie gibt, von vornherein stärker zu berücksichtigen. Sie machen damit die Domäne 3 (Methodologische Exaktheit der Leitlinienentwicklung) zu Ihrem Filter/Knock-Out-Kriterium. Ferner beschließen Sie, in Zukunft vermehrt systematisch nach Leitlinien zu suchen und kontroverse Fragen leitliniengestützt zu diskutieren.

15.9 Zusammenfassung

In Leitlinien wird Wissen zusammengefasst und bewertet, um es für die Anwendung in der täglichen Praxis zur Verfügung zu stellen. Für Leitlinien besteht, wie für andere Informationsquellen auch, die Notwendigkeit, diese hinsichtlich ihrer Validität, Relevanz und Anwendbarkeit zu beurteilen.

Hierfür steht mit DELBI (Deutsches Leitlinien-Bewertungs-Instrument) ein auf internationaler Methodik aufbauendes und an den Bedarf des deutschen Gesundheitssystems angepasstes Instrument zur Verfügung. Mit einer vergleichenden Bewertung mehrerer Leitlinien zum gleichen Thema gelingt ein Überblick über Stärken und Schwächen einzelner Leitlinien. Damit wird die Auswahl der für die eigenen Belange am besten geeigneten Leitlinie erleichtert.

Mit Checklisten zur Bewertung der formalen und methodischen Qualität lässt sich die inhaltliche Qualität von Leitlinien nicht unmittelbar erfassen. Die Beurteilung der Methodik ermöglicht jedoch, entsprechend dem Critical Appraisal anderer Quellen, eine Einschätzung darüber, ob systematische Methoden bei der Leitlinienformulierung angewandt wur-

den, um so das Risiko einer systematischen Verzerrung bei der Formulierung der Leitlinienempfehlungen möglichst gering zu halten.

Checklisten helfen dem Leitlinienanwender, subjektive und implizite Kriterien, mit denen jeder für sich die Qualität einer Informationsquelle einschätzt, durch explizite Kriterien zu ergänzen und damit möglichst nachvollziehbar zu machen.

Wie im Vorwort von DELBI dargestellt, vertreten wir die Auffassung, dass es derzeit wahrscheinlich keine Leitlinie gibt, die bei einer formalen und methodischen Bewertung mittels DELBI in jeder Domäne die maximal erreichbare Punktzahl erzielen würde. Eine „perfekte“ Leitlinie kann es in dieser Hinsicht möglicherweise niemals geben. Damit soll einerseits nicht ausgesagt werden, dass die gegenwärtig verfügbaren Leitlinien untauglich wären für den praktischen Einsatz. Andererseits soll aber auch nicht der Eindruck erweckt werden, dass es sich nicht lohne, kontinuierlich an der Verbesserung von Leitlinien zu arbeiten. DELBI sollte als Anreiz und Herausforderung begriffen werden, nicht als starre Vorgabe.

15.10 Ausblick: Nationale und internationale Aktivitäten zur Förderung der Leitlinienqualität

In den vergangenen Jahren hat das Interesse an der Entwicklung, Bewertung und Implementierung evidenzbasierter medizinischer Leitlinien weltweit zugenommen. Mittlerweile arbeiten zahlreiche nationale Institutionen sowie Wissenschaftlicher und Leistungserbringer im Gesundheitswesen in supranationalen Netzwerken auf dem Gebiet der Förderung von Leitlinien-Qualität und Nutzung [16; 17].

So hat auf beispielsweise der Europarat im Jahr 2001 unter Beteiligung von Deutschland, Österreich und der Schweiz Empfehlungen

für seine Mitgliedsstaaten zur Leitlinien-Methodik verabschiedet [13] [14]. Eine durch das Biomed 2-Programm der Europäischen Union geförderte Forschergruppe (die so genannte AGREE Collaboration) entwickelte das AGREE-Instrument zur formalen Qualitätsbewertung von Leitlinien, welches international großen Anklang gefunden hat [8]. Die AWMF [18] und die Spitzenverbände der Selbstverwaltung im deutschen Gesundheitswesen [19] etablierten – unter Nutzung vergleichbarer Instrumente – zunächst organisatorisch getrennte Programme zur Qualitätsförderung von Leitlinien. Methoden und Zielsetzungen dieser Projekte wurden 2001 zunächst durch das Leitlinienmanual vom AWMF und ÄZQ [20], seit 2003 durch das Programm für Nationale VersorgungsLeitlinien [21] und 2005 durch das DELBI-Instrument [1] harmonisiert.

Allerdings existierte bis 2002 – im Gegensatz zu den internationalen Kooperationen in den Bereichen HTA (siehe Kapitel 16) und Systematische Reviews (siehe Kapitel 43) kein institutionalisiertes supranationales Kommunikationsforum für Leitlinien-Institutionen und -Experten. Infolgedessen existierte weltweit über lange Jahre Parallelarbeit bei der Entwicklung von Leitliniendokumenten und Implementierungsstrategien. Um hier Abhilfe zu schaffen, vor allem um die Leitlinienarbeit inhaltlich und finanziell zu rationalisieren, wurde 2002 auf Initiative des ÄZQ hin das internationale Leitlinien-Netzwerk G-I-N (Guidelines International Network, <http://www.g-i-n.net>) gegründet [17].

G-I-N bemüht sich um Qualitätsentwicklung der Gesundheitsversorgung durch Unterstützung internationaler Zusammenarbeit auf dem Gebiet der systematischen Entwicklung von evidenzbasierten Leitlinien und ihrer Anwendung in Medizin und anderen Bereichen des Gesundheitswesens.

Die betrifft insbesondere:

- ▲ Förderung von Informationsaustausch, Ausbildung, Wissenstransfer und Zusammenarbeit zwischen Leitlinien-Program-

- men zur Berücksichtigung internationaler Standards;
- ▲ Verbesserung und Harmonisierung der Methodik zur systematischen Leitlinien-Entwicklung bei existierenden und neuen Leitlinienprogrammen;
 - ▲ Verbesserung der Methodik zur Verbreitung, Implementierung und Evaluation medizinischer Leitlinien;
 - ▲ Identifizierung von Prioritäten sowie Unterstützung von Forschung und Forschungstransfer auf dem Gebiet der Entwicklung, Verbreitung, Implementierung und Evaluation von Leitlinien;
 - ▲ Vernetzung von Organisationen zur verbesserten Koordination mit anderen Qualitätsinitiativen im Gesundheitswesen.

Wesentliche Instrumente zur Erreichung dieser Ziele sind – neben der Organisation internationaler Kongresse und Workshops – das G-I-N-Internet-Forum (<http://www.g-i-n.net>) mit der weltweit umfassendsten Leitlinien-Datenbank „International Guideline Library“. Sie umfasst Anfang 2007 mehr als 4000 Leitlinien und Leitlinien-Dokumente (Methodik-Manuale, Evidenztabelle, Implementierungshilfen und mehr).

Das Netzwerk entwickelte sich bis Anfang 2007 zu einer nahezu ausschließlich von nationalen Qualitätsagenturen getragenen Organisation (s. Tab. 15.5). Künftig wird zusätzlich der Einzelmitgliedschaft von Experten größere Bedeutung zugemessen.

Tab. 15.5: Institutionelle Mitglieder des Guidelines International Network, Stand Februar 2007.
Quelle: www.g-i-n.net

Australia: Joanna Briggs Institut. (JBI), Nat. Heart Found. of Australia (NHFA), Nat. Health & Medical Research Council (NHMRC), Nat. Inst. for Clinical Studies (NICS), Therapeutic Guidelines Ltd (TGL)
Austria: Ärztekammer Wien (AEKW)
Belgium: Belgian Centre for Evidence-Based Medicine (CEBAM); Domus Medica (DM)
Brazil: Brazilian Medical Association (AMB)
Canada: Québec Govern. Agency for Health Serv. & TA ; Guidelines Advisory Committee, Ontario
Denmark: Danish National Board of Health (SST)
Estonia: Estonian Health Insurance Fund (EHIF)
Finland: Fin. Med. Society (DUODECIM); Fin. Office for HTA (FinOHTA), Centre for Pharmacother. Development (ROHTO)
France: French National Health Authority (HAS); National Federation of Cancer Centres (FNCLCC)
Germany: Ärztekammer Berlin, ÄZQ, AWMF, BQS, G-BA, IQWiG, Zahnärztl. Zentralstelle QS (ZZQ)
Hong Kong: Hospital Authority Hong Kong Special Admin. Region (HA), Profess. Devel. & Quality Assurance, Department of Health (PDQA-DoH)
Ireland: The Mental Health Commission (MHC)
Italy: Regional Health Agency Emilia Romagna (ASR); Italian EbM Group (GIMBE)
Japan: Med. Inform. Network Distrib. Service Center, Japan Council for Quality Health Care
Malaysia: HTA Unit, Ministry of Health
Mexico: National Center for Health Technology Excellence
Netherlands: Dutch Association of Comprehensive Cancer Centres (ACCC), Dutch Institute for Healthcare Improvement (CBO), Royal Dutch Society for Physical Therapy (KNGF), Dutch College of General Practitioners (NHG); Trimbos-Inst. - NL Institute of Mental Health & Addiction

Tab. 15.5: Fortsetzung

New Zealand: NZ Accident Compensation Corporation, NZ Guidelines Group (NZGG)
Norway: Norwegian Directorate for Health and Social Affairs
Portugal: Center for EbM, Univ. of Lisbon School of Medicine, Institute for Quality in Healthcare (IQS)
Romania: Nat. Center for Studies in Family Medicine (CNSMF), Center for Health Policies and Services (CPSS), National School of Public Health and Health Services Management (INCDIS)
Slovenia: Slovene Guidelines Group (SGG)
Spain: Catalan Agency for HTA and Research (AATRM), Basque Office for HTA (OSTEBA), Spanish Network for Research on Guidelines (REDEGUIAS)
Sweden: National Board of Health and Welfare (Socialstyrelsen)
Switzerland: Swiss Federal Office of Public Health (BAG); Clinical Epidemiology Centre, University Hospital Lausanne (CepiC); Swiss Medical Association (FMH)
Taiwan: Center for Health Policy Research & Development, Nat. Health Research Institutes
United Kingdom: National Institute for Health and Clinical Excellence (NICE), Royal College of Nursing (RCN); Sowerby Centre for Health Informatics at Newcastle (SCHIN); Scottish Intercollegiate Guidelines Network (SIGN)
USA: Agency for Health Research and Quality (AHRQ), Center for International Rehabilitation (CIR); National Kidney Foundation (NKF)
International: The AGREE Research Trust (ART), European Union of Medical Specialists (UEMS), European Region of the World Confederation of Physical Therapy (WCPT), World Health Organisation (WHO), The World Medical Association (WMA)

Tab. 15.6: Statements in DELBI

Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI)					
Fassung Juni 2005 – Kurzversion zu Informationszwecken, Teil 1 von 2 –					
Zur Bewertung NUR die ausführliche Langfassung verwenden!					
	Domäne 1: Geltungsbereich und Zweck	1	2	3	4
1	Das Gesamtziel der Leitlinie ist differenziert beschrieben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	Die in der Leitlinie behandelten medizinischen Fragen/Probleme sind differenziert beschrieben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Die Patienten, für die die Leitlinie gelten soll, sind eindeutig beschrieben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Domäne 2: Beteiligung von Interessengruppen	1	2	3	4
4	Die Entwicklergruppe der Leitlinie schließt Mitglieder aller relevanten Berufsgruppen ein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Die Ansichten und Präferenzen der Patienten wurden ermittelt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	Die Anwenderzielgruppe der Leitlinie ist definiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	Die Leitlinie wurde in einer Pilotstudie von Mitgliedern der Anwenderzielgruppe getestet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Bewertung 1: Trifft überhaupt nicht zu, Bewertung 4: Trifft uneingeschränkt zu

Tab. 15.6: Fortsetzung

Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI) Fassung Juni 2005 – Kurzversion zu Informationszwecken, Teil 2 von 2 – Zur Bewertung NUR die ausführliche Langfassung verwenden!					
	Domäne 3: Methodologische Exaktheit der Leitlinien-Entwicklung	1	2	3	4
8	Bei der Suche nach der Evidenz wurden systematische Methoden angewandt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	Die Kriterien für die Auswahl der Evidenz sind klar beschrieben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	Die zur Formulierung der Empfehlungen verwendeten Methoden sind klar beschrieben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11	Bei der Formulierung der Empfehlungen wurden gesundheitlicher Nutzen, Nebenwirkungen und Risiken berücksichtigt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	Die Verbindung zwischen Empfehlungen und der zugrunde liegenden Evidenz ist explizit dargestellt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13	Die Leitlinie ist vor ihrer Veröffentlichung durch externe Experten begutachtet worden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	Ein Verfahren zur Aktualisierung der Leitlinie ist angegeben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Domäne 4: Klarheit und Gestaltung	1	2	3	4
15	Die Empfehlungen der Leitlinie sind spezifisch und eindeutig.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16	Die verschiedenen Handlungsoptionen für das Versorgungsproblem sind dargestellt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17	Schlüsselempfehlungen der Leitlinie sind leicht zu identifizieren.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18	Es existieren Instrumente bzw. Materialien, die die Anwendung der Leitlinie unterstützen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Domäne 5: Generelle Anwendbarkeit	1	2	3	4
19	Die möglichen organisatorischen Barrieren gegenüber der Anwendung der Empfehlungen werden diskutiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	Die durch die Anwendung der Empfehlungen der Leitlinie möglicherweise entstehenden finanziellen Auswirkungen werden berücksichtigt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21	Die Leitlinie benennt wesentliche Messgrößen für das Monitoring und/oder die Überprüfungskriterien.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Domäne 6: Redaktionelle Unabhängigkeit	1	2	3	4
22	Die Leitlinie ist redaktionell von der (den) finanzierenden Organisation(en) unabhängig.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23	Interessenkonflikte von Mitgliedern der Leitlinienentwicklungsgruppe wurden dokumentiert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Bewertung 1: Trifft überhaupt nicht zu, Bewertung 4: Trifft uneingeschränkt zu

Tab. 15.6: Fortsetzung

Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI) Fassung Juni 2005 – Kurzversion zu Informationszwecken, Teil 2 von 2 – Zur Bewertung NUR die ausführliche Langfassung verwenden!					
Domäne 7: Anwendbarkeit im deutschen Gesundheitssystem		1	2	3	4
24	Es liegen Empfehlungen zu präventiven, diagnostischen, therapeutischen und rehabilitativen Maßnahmen in den verschiedenen Versorgungsbereichen vor.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25	Es existieren Angaben, welche Maßnahmen unzweckmäßig, überflüssig oder obsolet erscheinen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26	Die klinische Information der Leitlinie ist so organisiert, dass der Ablauf des medizinischen Entscheidungsprozesses systematisch nachvollzogen wird und schnell erfassbar ist.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27	Es ist eine Strategie/ein Konzept für die einfache Zugänglichkeit und für die Verbreitung der Leitlinie dargelegt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28	Ein Konzept zur Implementierung der Leitlinie wird beschrieben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29	Der Leitlinie ist eine Beschreibung zum methodischen Vorgehen (Leitlinien-Report) hinterlegt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Bewertung 1: Trifft überhaupt nicht zu, Bewertung 4: Trifft uneingeschränkt zu

Literatur

- [1] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI). Fassung 2005/2006. Z Arztl Fortbild Qualitätssich. (2005), 99, 468–519
- [2] Grimshaw JM et al., Changing physicians' behavior: what works and thoughts on getting more things to work. J Contin Educ Health Prof. (2002), 22, 237–243
- [3] Grol R et al., Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. BMJ. (1998), 317, 858–861
- [4] Leape LL et al., Adherence to practice guidelines: the role of specialty society guidelines. Am Heart J. (2003), 145, 19–26
- [5] Field MJ et al. (1990) Clinical Practice Guidelines: Directions for a New Program. National Academy Press, Washington DC
- [6] Field MJ, Lohr KN (1992) Guidelines for clinical practice: from development to use. Institute of Medicine, Washington DC
- [7] Ollenschläger G, Thomeczek C, Ärztliche Leitlinien-Definitionen, Ziele, Implementierung. Z Arztl Fortbild Qualitätssich. (1996), 90, 355–361
- [8] The AGREE Collaboration, Ärztliche Zentralstelle Qualitätssicherung (ÄZQ), Verbindung der Schweizer Ärztinnen und Ärzte FMH. Checkliste zur Qualitätsbeurteilung von Leitlinien. AGREE Instrument – Deutschsprachige Version. <http://www.agreecollaboration.org/pdf/de.pdf> (05.10.05)
- [9] Burgers JS, Guideline quality and guideline content: are they related? Clin Chem. (2006), 52, 3–4
- [10] Mant J, McManus RJ, Hare R, Applicability to primary care of national clinical guidelines on blood pressure lowering for people with stroke: cross sectional study. BMJ. (2006), 332, 635–637
- [11] van Weel C, Schellevis FG, Comorbidity and guidelines: conflicting interests. Lancet. (2006), 367, 550–551
- [12] Watine J et al., Conflict between guideline methodologic quality and recommendation validity: a potential problem for practitioners. Clin Chem. (2006), 52, 65–72
- [13] Europarat, Entwicklung einer Methodik für die Ausarbeitung von Leitlinien für optimale medizinische Praxis. Empfehlung Rec (2001)13 des Europarates und Erläu-

- terndes Memorandum. Deutschsprachige Ausgabe. Z Arztl Fortbild Qualitätssich. (2002), 96 Suppl III, 1–60
- [14] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Das Deutsche Leitlinien-Clearingverfahren 1999–2005 – Hintergrund, Zielsetzung, Ergebnisse. Abschlussbericht. <http://www.leitlinien.de/clearingverfahren/index/clearing/view> (28.06.06)
- [15] Ärztliche Zentralstelle Qualitätssicherung (ÄZQ), Checkliste „Methodische Qualität von Leitlinien“. Dt Arztebl. (2000), 97, A 1170–2
- [16] Burgers J, Grol R, Klazinga N, van der Bij A, Mäkelä M, Zaat J, Internationaler Vergleich von 19 Leitlinien-Programmen-Eine Übersicht der AGREE Collaboration. Z Arztl Fortbild Qualitätssich. (2003), 97, 81–88
- [17] Ollenschläger G, Marshall C, Qureshi S, Rosenbrand K, Burgers J, Mäkelä M, Slutsky J, Improving the quality of health care: using international collaboration to inform guideline programmes by founding the Guidelines International Network (G-I-N). Qual Saf Health Care. (2004), 13, 455–460
- [18] Kopp I, Encke A, Lorenz W, Leitlinien als Instrument der Qualitätssicherung in der Medizin. Das Leitlinienprogramm der Arbeitsgemeinschaft Wissenschaftlicher Medizinischer Fachgesellschaften (AWMF). Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz. (2002), 45, 223–233
- [19] Ollenschläger G, Thomeczek C, Bungart B, Lampert U, Arndt S, Kolkman FW, Oesingmann U, Das Leitlinien-Clearingprogramm der Selbstverwaltungskörperschaften im Gesundheitswesen. Ein Projekt zur Qualitätsförderung in der Medizin. Gesundheitswesen. (1999), 61, 105–111
- [20] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Ärztliche Zentralstelle Qualitätssicherung (ÄZQ), Das Leitlinien-Manual von AWMF und ÄZQ. Z Arztl Fortbild Qualitätssich. (2001), 95 Suppl 1, 1–84
- [21] Ollenschläger G, Kopp I, Lelgemann M, Sängler S, Heymans L, Thole H, Trapp H, Lorenz W, Selbmann HK, Encke A, Nationale Versorgungsleitlinien von BÄK, AWMF und KBV: Hintergrund, Methodik und Instrumente. Med Klin (Munich). (2006), 101, 840–845

16 Kritische Bewertung von Health Technology Assessment-Berichten

Matthias Perleth, Dagmar Lühmann

16.1 Einleitung

16.1.1 Health Technology Assessment (HTA) – Ursprünge und Definition

Die Wurzeln von HTA liegen in den Arbeiten des „Office of Technology Assessment“ (OTA) in den USA. Dort wurden seit Mitte der 1960er-Jahre multidisziplinäre Technikfolgen-Abschätzungen (TA) vorgenommen und zwar für solche Technologien, für die gesamtgesellschaftliche Auswirkungen anzunehmen waren (z.B. Umwelt-, Verkehrs-, Weltraum-, Informationstechnologien). Mit der Einführung eines Gesundheitsprogramms 1976 wurde die Grundlage des eigentlichen „Health Technology Assessment“ (HTA) gelegt.

HTA hat sich seitdem in zwei Richtungen weiter entwickelt, die zwar häufig die gleichen Themen, diese aber aus unterschiedlichen Perspektiven bearbeiten. Sie sind somit komplementär zu sehen: die parlamentarische Technikfolgen-Abschätzung mit direkt an die Legislative gerichteten Bewertungen und die pragmatische Technologiebewertung, das eigentliche HTA. Im Gegensatz zum parlamentarischen TA sind die Adressaten für HTA Entscheidungsträger auf unterschiedlichen Ebenen innerhalb der Gesundheitssysteme. HTA-Berichte fokussieren vor allem auf klinische und ökonomische Technikfolgen. Soziale, rechtliche und ethische Implikationen gehören zwar zu den Anforderungen an ein vollständiges Assessment, nehmen aber in der Praxis einen eher untergeordneten Stellenwert ein.

Health Technology Assessment (HTA) [...] ist eine Form der Politikfeldanalyse (Policy Research) die systematisch kurz- und langfristige Konsequenzen der Anwendung einer medizinischen Technologie, einer Gruppe verwandter Technologien oder eines technologiebezogenen Sachverhalts untersucht. Das Ziel von HTA ist die Unterstützung von Entscheidungen in Politik und Praxis. Grundlegerend für HTA ist die Ausrichtung auf Entscheidungsfindung sowie der multidisziplinäre und umfassende Ansatz [...] [1].

HTA betrachtet medizinische Technologien umfassend und führt Analysen aus unterschiedlichen Perspektiven durch. Darin sind enthalten: Studien zu ethischen und sozialen Auswirkungen von Technologien; Faktoren, die die Diffusion medizinischer Technologien hemmen oder beschleunigen; Untersuchung der Effekte, Regelungen zur Diffusion und Nutzung (Qualitätssicherung) von Technologien und Vorschläge zur Änderung solcher Regelungen; Studien zu Variationen in der Nutzung von Technologien. Der elementarste Bestandteil von HTA besteht darin, Nutzen und Kosten einer Technologie zu bestimmen und somit zu einer wirtschaftlichen Verwendung von Ressourcen im Gesundheitswesen beizutragen.

In diesem umfassenden Kontext definiert sich HTA nicht durch eine Reihe von Methoden, sondern durch seine Intention bzw. Funktion. Die Überprüfung der technischen Eigenschaften eines Arzneimittels oder eines Medizinproduktes als Bestandteil einer Regu-

lationsentscheidung kann ebenso als HTA bezeichnet werden wie eine ethische Analyse von gentherapeutischen Verfahren, die einer Kostenübernahme-Entscheidung vorausgeht. Ungeachtet dieser politischen Zielsetzung muss HTA auf wissenschaftlicher Grundlage durchgeführt werden und sich wissenschaftlicher Methoden bedienen. Der

Prozess der Technologiebewertung muss integer durchgeführt werden und die Ergebnisse müssen glaubwürdig und nachvollziehbar sein. Die häufigste Aktivität von HTA ist die systematische Recherche und Zusammenstellung der verfügbaren Informationen zu einer medizinischen Technologie, insbesondere zur Wirksamkeit und Kosten-Wirksam-

Tab. 16.1: Komponenten und Methodenspektrum von HTA

Baustein des HTA	Bedeutung	Methoden
Statusbestimmung hinsichtlich: Regulation/Zulassung Kostenerstattung Diffusion und Nutzungshäufigkeit	Überblick über den gegenwärtigen Status einer Technologie hinsichtlich rechtlicher und versorgungspraktischer Aspekte, auch international vergleichend	Analyse von Dokumenten und Verordnungen; Umfragen bei zuständigen Organisationen (national und/oder international)
Bewertung des klinischen Nutzens bzw. des Nutzens für Patienten und der Risiken	Systematische Darstellung der Wirkungen von Technologien auf den Gesundheitszustand und die Lebensqualität sowie der unerwünschten Effekte	Systematische Übersichten und Meta-Analysen zu Nutzen und Risiken diagnostischer und therapeutischer Technologien; Durchführung von klinischen Studien
Bewertung der Wirtschaftlichkeit: nicht-vergleichend/vergleichend	Analyse der ökonomischen Effekte von medizinischen Technologien, inklusive Ermittlung der vergleichenden Effizienz und der Lebensqualität	Systematische Übersichten und Entscheidungsanalysen gesundheitsökonomischer Studien, Durchführung von gesundheitsökonomischen Primärstudien
Fallstudien	Klärung der Bedingungsfaktoren wichtiger Aspekte der Ausbreitung und Nutzung von (paradigmatischen) Technologien	Tiefgehende Analyse einzelner Technologien in ihrem politischen, organisatorischen und finanziellen Kontext
Innovationsbezogenes HTA	Begleitende Evaluation von Technologien in der Entwicklungsphase vor Markteinführung	Klinische Studien, epidemiologische und ökonomische Analysen, Surveys, Delphi-Panels, Modellierungen
Implikationen für die Organisation	Einschätzung des Einflusses der Einführung und Anwendung von Technologien auf die Organisation der Gesundheitsversorgung	Analyse struktureller und organisatorischer Rahmenbedingungen und Abschätzung der Einflüsse der Technologie auf Finanz- und Patientenströme sowie auf Anbieterstrukturen
Soziale, gesellschaftliche, psychologische und ethische Implikationen	Analyse der mit der Nutzung einer Technologie einhergehenden ethischen und sozialen Probleme	Qualitative Studien auf der Basis von Literaturrecherchen und Umfragen, Interviews; Auswertung von Dokumenten

Quelle: eigene Zusammenstellung

keit. Ziel ist die Unterstützung verschiedener Entscheidungen, die auch die Generierung von neuer Evidenz betreffen können, wie z.B. die Durchführung von randomisierten kontrollierten Studien.

Trotz dieses umfassenden Spektrums ist HTA keine eigenständige Disziplin. Tatsächlich handelt es sich um einen systematischen, interdisziplinären Prozess, der auf die Zusammenführung und Kondensierung von wissenschaftlicher Evidenz mit anderen Informationen zielt. HTA involviert u.a. Kliniker, Ökonomen, Sozialwissenschaftler, Naturwissenschaftler aus den Bereichen Public Health und Versorgungsforschung, Ingenieure und Ethiker. Außerdem werden zunehmend Patientinnen, die Öffentlichkeit bzw. ihre Repräsentanten in den HTA-Prozess einbezogen.

Dem breiten und interdisziplinären Ansatz entsprechend, werden je nach Fragestellung unterschiedliche Facetten einer Technologie untersucht. Ein Überblick über die Komponenten und häufig angewandte Methoden findet sich in Tabelle 16.1.

16.2 Szenario

Sie leiten in einer Krankenkasse eine Arbeitsgruppe, die Verträge mit Leistungserbringern erarbeitet. Sie sollen nun einen Vertrag zur Versorgung von Patientinnen mit Brustkrebs mit dem neuen Medikament Trastuzumab in spezialisierten Zentren ausarbeiten. Bisher wurde das Medikament bei Frauen mit metastasiertem Brustkrebs eingesetzt. Nun stellt sich die Frage, ob auch ein Einsatz bei anderen Indikationen infrage kommt, insbesondere bei Brustkrebs im Frühstadium. Sie wollen außerdem wissen, welche Kosten-Nutzen-Relation mit der Therapie verbunden ist. Sowohl Patientenorganisationen wie auch Fachgesellschaften und Industrievertreter sind in den Medien präsent und fordern einen uneingeschränkten Einsatz von Trastuzumab bei Brustkrebs.

16.2.1 Hintergrund

Trastuzumab ist ein monoklonaler Antikörper, der das Protein HER-2/neu (humaner epidermaler Wachstumsfaktor-Rezeptor) blockiert, das auf der Oberfläche bestimmter Brustkrebszellen überexprimiert wird. HER-2/neu erhöht die Sensitivität für Wachstumsfaktoren. Die HER-2-positiven Tumore weisen dadurch ein besonders aggressives Wachstum auf. Durch die Blockade des Rezeptors wird das Wachstum dieser Zellen gehemmt. Für die Therapie kommen etwa 20–30% der Betroffenen infrage, bei denen >10% der Brustkrebszellen eine Überexpression aufweisen. Deshalb ist vor der Therapie ein entsprechender immunhistochemischer Test an Biopsiegewebe durchzuführen. Trastuzumab wird als Ergänzung zur Chemotherapie eingesetzt. Trastuzumab ist allerdings potenziell kardiotoxisch, da auch Herzmuskelzellen HER-2-Rezeptoren aufweisen. Seit 2005 wird über Studien berichtet, die den Einsatz von Trastuzumab bei frühen Brustkrebsstadien untersuchen. Im Mai 2006 wurde Trastuzumab auch für diese Indikation von der Europäischen Arzneimittelbehörde EMA zugelassen [2]. Das führt zu einem möglichen Dilemma bei Patientinnen mit Brustkrebs im Frühstadium, die unter kurativer Zielsetzung behandelt werden. Hier muss nun der mögliche Nutzen der Therapie mit Trastuzumab mit einer möglichen Herzschildigung und anderen Nebenwirkungen abgewogen werden.

16.2.2 Die Evidenz suchen

Eine orientierende Literaturrecherche in MEDLINE unter dem Stichwort „Trastuzumab“ ergibt über 1500 Treffer, davon über 500 Reviewartikel. Zusätzlich recherchieren Sie in einer internationalen HTA-Datenbank [http://www.york.ac.uk/inst/crd/crddatabases.htm] nach HTA-Berichten, die auch eine

ökonomische Evaluation enthalten. Sie erhalten unter dem Stichwort „Trastuzumab“ 22 Treffer, darunter auch einen aktuellen Bericht des belgischen HTA-Instituts KCE, das als nachgeordnete Einrichtung des Gesundheitsministeriums arbeitet (<http://www.kenniscentrum.fgov.be>). Dieser scheint Ihre Anforderungen zu erfüllen [3]. Sie laden den vollständigen Bericht aus dem Internet herunter und beschließen, ihn nach einer ersten Durchsicht auszuwerten.

16.2.3 Qualitätsmerkmale von HTA-Berichten

HTA ist in vielen Gesundheitssystemen mittlerweile eine wichtige Hilfe bei Entscheidungen zur Kostenübernahme, Finanzierung von oder Investitionen in medizinische Technologien geworden. Damit die Gefahr von Fehlentscheidungen so gering wie möglich gehalten wird, sollten HTAs hohen methodischen Anforderungen genügen, insbesondere bei der Bewertung der klinischen Wirksamkeit und der ökonomischen Effekte, da hierdurch am ehesten Schaden (oder entgangener Nutzen) für Patienten und das Gesundheitswesen insgesamt resultieren kann. Ein weiterer wichtiger Grund für eine hohe Qualität ist die zunehmende internationale Zusammenarbeit von HTA-Einrichtungen, die sich beispielsweise im Austausch von HTA-Berichten ausdrückt. Dabei liegt es in der Natur von HTA, dass eine Bewertung, die unter Entscheidungsdruck in einem spezifischen regionalen oder nationalen Kontext erstellt wurde, zu Schlussfolgerungen kommt, die nicht unbedingt auf andere Systeme übertragbar sind. Ein internationaler Austausch von HTA Informationen kann nur dann gelingen, wenn in den Berichten die Komponenten Assessment und Appraisal eindeutig voneinander zu trennen sind. Im

Gegensatz zum Begriff „Critical Appraisal“ in der EbM meint Appraisal hier die Bewertung der Datenlage im Systemkontext; Assessment bezeichnet dagegen in der HTA-Nomenklatur die Sammlung, Aufbereitung und Kenntnisnahme der wissenschaftlichen Datenlage. Diese Begriffstrennung hat sich zunächst in England mit der Gründung des National Institute for Clinical Excellence (NICE) etabliert und ist nun international anerkannt, da die Sammlung und Aufbereitung der Evidenz zumeist getrennt von der Umsetzung evidenzbasierter Empfehlungen im Gesundheitswesen (die ja noch von vielen anderen Faktoren abhängt) stattfindet [4; 5]. Dies spiegelt sich auch in der Praxis des Gemeinsamen Bundesausschusses wider.

HTA-Berichte müssen daher einer Reihe von Anforderungen hinsichtlich des Erstellungsprozesses und des Berichtsformats genügen. Eine Voraussetzung für die Einschätzung der Qualität von HTA-Berichten ist eine transparente und nachvollziehbare Darstellung der Methoden, mit denen Schlussfolgerungen bzw. Empfehlungen erreicht wurden. Dies sollte anhand einer ausführlichen, dem Hauptbericht vorangestellten Zusammenfassung (Executive Summary, Scientific summary report) möglich sein, insbesondere wenn auf HTA-Berichte anderer Agenturen zurückgegriffen werden soll [6].

Fast jede HTA-Einrichtung hat ein eigenes Format und eigene Arbeitsmethoden entwickelt. Dies trägt der Tatsache Rechnung, dass HTA immer in ein regionales Entscheidungsgefüge mit jeweils unterschiedlichen Anforderungen eingebunden ist, was sich auch im Berichtsformat widerspiegelt. Dies enthebt die Forderung nach größtmöglicher Transparenz allerdings nicht ihrer Gültigkeit. Es hat also seine Berechtigung, aus den Berichtsteilen, die ein HTA mindestens enthalten sollte, Fragen nach ihrer Qualität abzuleiten (s. Tab. 16.2).

Tab. 16.2: Kriterien für die Beurteilung der Qualität von HTA-Berichten [7]

Kriterium	Checkliste
Darstellung des Hintergrunds der Untersuchung (policy question)	<p>Wurde angegeben, warum das HTA durchgeführt wurde?</p> <p>Gibt es eine Begründung, warum das HTA zum jetzigen Zeitpunkt durchgeführt wurde (z.B. umstrittene Indikationsausweitung)?</p> <p>Wurde angegeben, welche Entscheidungen durch das HTA unterstützt werden sollen/können?</p> <p>Wurde angegeben, wer das HTA angefordert bzw. in Auftrag gegeben und finanziert hat?</p>
Formulierung von spezifischen Fragestellungen	<p>Wurden Zielkondition, Patientencharakteristika (z.B. Krankheitsstadien, Schweregrad, Geschlecht, Alter), relevante Interventionen bzw. Vergleiche zwischen Interventionen sowie (patienten)relevante Endpunkte festgelegt?</p>
Technische Charakteristika der Technologie	<p>Wurde beschrieben, inwieweit technische Merkmale einer Technologie direkt mit der Wirksamkeit verknüpft sind (z.B. bei bildgebenden Verfahren)?</p> <p>Wurden weitere Einflussfaktoren beschrieben (z.B. Qualifikation des Personals, technische Qualitätssicherung, Risiken für Patienten, Personal, Umwelt)?</p>
Angaben zum Status quo der Technologie	<p>Wurden Nutzungsmuster bzw. Diffusion (z.B. Grad der Verbreitung außerhalb von Universitätskliniken), Indikationsspektrum und zeitliche Entwicklungstrends (z.B. Änderung in der Nutzungshäufigkeit) beschrieben?</p> <p>Wurde der rechtliche (z.B. Marktzulassung) Status der Technologie analysiert?</p>
Systematische Evaluation der Sicherheit und klinischen Wirksamkeit	<p>Wurde ein Protokoll erstellt?</p> <p>Wurde die Literaturrecherche systematisch durchgeführt und dokumentiert (inklusive Suchstrategien, Datenquellen, Jahrgänge)?</p> <p>Wurden Ein- und Ausschlusskriterien für Primärstudien festgelegt?</p> <p>Wurde eine Qualitätsbewertung der einzuschließenden Studien durchgeführt?</p> <p>Wurde die Extraktion der Daten aus den eingeschlossenen Studien beschrieben?</p> <p>Wurde die Auswertungsstrategie (z.B. qualitative tabellarische Übersicht, Meta-Analyse) nachvollziehbar beschrieben?</p> <p>Sind die Ergebnisse detailliert dokumentiert (z.B. Übersichtstabellen)?</p> <p>Sind die Schlussfolgerungen kompatibel mit den Ergebnissen?</p>
Gesundheitsökonomische Evaluation	<p>Ist die Methodik (analog der Übersicht über die klinische Wirksamkeit) nachvollziehbar dargestellt?</p> <p>Ist die Perspektive der gesundheitsökonomischen Bewertung (z.B. Krankenkassen, Gesellschaft, Patienten) definiert?</p> <p>Ist die Übertragbarkeit (z.B. Mengengerüste, Kaufkraftparitäten und systemabhängige Variablen) ausreichend diskutiert?</p> <p>Wurden Annahmen und Voraussetzungen ausreichend begründet (z.B. Diskontierungsraten, Sensitivitätsanalysen)?</p>
Evaluation der Lebensqualität	<p>Wurden für Patienten relevante Aspekte der Lebensqualität berücksichtigt?</p> <p>Wurden validierte Instrumente zur Messung der Lebensqualität verwendet?</p>

Tab. 16.2: Fortsetzung

Kriterium	Checkliste
Entscheidungsanalyse	<p>Wurden alle wichtigen Strategien und Endpunkte berücksichtigt?</p> <p>Wurden explizite und ausreichend empfindliche Methoden zur Identifikation, Auswahl und Zusammenfassung der Evidenz in Wahrscheinlichkeiten verwendet?</p> <p>Wurden die Quellen der verwendeten Variablen benannt?</p> <p>Wurde die Zuverlässigkeit der verwendeten Variablen angegeben (z.B. Herkunft aus Studien oder Statistiken, Schätzung, Annahme, Extrapolation)?</p> <p>Wie ist die Güte der Evidenz, die in die Analyse eingegangen ist?</p> <p>Wurden begründete Sensitivitätsanalysen durchgeführt?</p>
Zusammenhang von Organisationsstrukturen und -abläufen mit der Technologie	<p>Wurden Aspekte der Organisationsstruktur diskutiert (z. B. Entscheidungskompetenzen [klinisch, Finanzierung], Personal [Bedarf, Qualifikationsprofil], administrative Zuordnung der Technologie, Anfälligkeit für interne [z.B. Personalmangel, technische Probleme] und externe [z.B. Budgetrestriktionen, erhöhte Nachfrage] Effekte, Datentransfer/Dokumentation [z.B. Erfassung der Teilnehmerate von Screening-Programmen])?</p>
Diskussion der Generalisierbarkeit/Übertragbarkeit	<p>Wurde die Generalisierbarkeit der Ergebnisse auf andere als die in Studien untersuchten Patientenpopulationen diskutiert (z.B. anderer Altersgruppe, Geschlecht, Krankheitsstadien)?</p> <p>Wurde die Übertragbarkeit der Ergebnisse auf andere Versorgungskontexte diskutiert (z.B. Epidemiologie, Diffusion, Versorgungs- und Organisationsstruktur, Vergütungssystem, Zugang, Patientenpräferenzen)?</p>
Einschätzung ethischer, sozialer und juristischer Implikationen	<p>Wurden die jeweiligen ethischen, sozialen, kulturellen, systembedingten und juristischen Besonderheiten im Zusammenhang mit der zur Diskussion stehenden Technologie berücksichtigt?</p>
Peer-Review, Revisionsdatum und Interessenskonflikte	<p>Wurden Angaben zum (internen/externen) Peer-Review gemacht?</p> <p>Ist ein Revisionsdatum angegeben bzw. sind Kriterien für eine notwendig werdende Revision angegeben (z.B. neue Studienergebnisse werden erwartet, die die Schlussfolgerungen/Empfehlungen beeinflussen könnten)?</p> <p>Liegen Erklärungen zu Interessenkonflikten von den am HTA beteiligten Personen vor?</p>

16.2.4 Die Evidenz bewerten

Wurde der Hintergrund der Untersuchung dargestellt?

Hintergrund des HTA-Berichts ist ein Auftrag des belgischen Gesundheitsministeriums an das KCE, die Wirksamkeit und Kosten (Budget-Impact) aus der Perspektive der Krankenkasse der Therapie mit Trastuzumab im Frühstadium des HER2-positiven Mammakarzinoms zu evaluieren. Das Ministerium reagiert damit auf Forderungen der Patientenvertreter, Trastuzumab für diese Indikation zu finanzieren.

Wurden spezifische Fragestellungen formuliert?

Ausgewertet werden Studien, die Trastuzumab in Kombination bzw. ergänzend zur Chemotherapie bei Frauen mit Brustkrebs im Frühstadium mit oder ohne Lymphknotenbefall hinsichtlich des krankheitsfreien Überlebens im Vergleich zur Standardtherapie testen. Zusätzlich werden Daten zum Schädigungspotenzial von Trastuzumab ausgewertet und es wird ein Sicherheitsprofil erstellt. Diese Daten zur Wirksamkeit und Sicherheit werden für die Ermittlung der Kos-

tenwirksamkeit mit Berechnung der inkrementellen Kosteneffektivitätsratio (ICER), d.h. der zusätzlichen Kosten für ein zusätzliches gewonnenes Lebensjahr, und zur Berechnung der finanziellen Gesamtbelastung des Gesundheitssystems vor dem Hintergrund der Epidemiologie des Mammakarzinoms in Belgien herangezogen.

Sind die technischen Charakteristika des Verfahrens beschrieben?

Das Medikament Trastuzumab wird ausführlich beschrieben, einschließlich Herstellung, Vertrieb und postuliertem Wirkmechanismus.

Wurden Angaben zum Status quo der Technologie gemacht?

Hierzu finden sich keine Angaben in dem Bericht.

Wurde eine systematische Evaluation der Sicherheit und klinischen Wirksamkeit durchgeführt?

Die Quellen für die Bewertung der Wirksamkeit und Sicherheit des HTA-Berichts sind eine nicht dokumentierte Literaturrecherche, eine Recherche im Studienregister clinicaltrials.gov und eine Anfrage beim Hersteller von Trastuzumab. Ein Dossier mit Ergebnissen auch unpublizierter Studienergebnisse, wie sie z.B. der EMEA vorgelegt werden, war den Autoren nicht zugänglich. Die Autoren schließen daher nicht aus, dass die im HTA-Bericht bewertete Studienlage unvollständig ist. Die Studienergebnisse werden tabellarisch-deskriptiv dargestellt, außerdem wird eine Meta-Analyse für die wichtigsten Endpunkte erstellt. Es wurden sowohl publizierte wie auch unpublizierte Studien ausgewertet.

Wurde eine gesundheitsökonomische Evaluation durchgeführt?

Die gesundheitsökonomische Bewertung stellt tatsächlich den Schwerpunkt des HTA-

Berichts dar. Die Vorgehensweise entspricht den belgischen Leitlinien für pharmakoökonomische Evaluationen, die ebenfalls vom KCE publiziert wurden. Im Gegensatz zur Bewertung der Wirksamkeit und Sicherheit werden die Quellen und Annahmen detailliert dargestellt. Zunächst werden die Kosten für die Restlebenserwartung modelliert. Dabei wird die Perspektive der Krankenversicherung eingenommen. Die Daten für die Modellierung werden aus Studien, belgischen Statistiken und einer Umfrage unter belgischen Krankenhäusern entnommen. Die Grenze der Vertretbarkeit wird arbiträr auf 30 000 Euro pro gewonnenes Lebensjahr festgelegt. Die inkrementellen Kosten setzen sich vor allem aus dem HER2-Test, den Kosten für Trastuzumab und der Bestimmung der linksventrikulären Ejektionsfraktion (LVEF) zusammen, außerdem aus Folgekosten für die Therapie eines Herzversagens durch Trastuzumab. Dem stehen die inkrementellen Effekte, insbesondere die Zeit bis zur Krankheitsprogression bzw. gewonnene Lebensjahre, gegenüber. Die Modellierung wird für verschiedene relevante Szenarien im Rahmen eines Markov-Modells erstellt, Diskontraten für Kosten und Effekte sind angegeben, außerdem wird eine probabilistische Sensitivitätsanalyse durchgeführt, um die Robustheit der Ergebnisse zu prüfen. Die Quellen für die verwendeten Variablen sowie der Grad der Unsicherheit (z.B. publiziert, arbiträr, geschätzt, kalkuliert) werden dargelegt. Schließlich werden die Gesamtkosten für das belgische Gesundheitssystem für die relevanten Szenarien berechnet.

Wurde der Einfluss auf die Lebensqualität evaluiert?

Da keine Daten zur Lebensqualität unter der Therapie mit Trastuzumab verfügbar waren, wird die Lebensqualität in der ökonomischen Evaluation nicht berücksichtigt. Eine Schwäche der Modellierung ist sicherlich die fehlende Berücksichtigung der Lebensquali-

tät, insbesondere wenn es infolge der Therapie mit Trastuzumab zu einer Herzinsuffizienz kommt.

Wurde eine Entscheidungsanalyse durchgeführt?

Es wird ein Markov-Modell konstruiert (Details s.o.).

Wurde der Zusammenhang von Organisationsstrukturen und -abläufen mit der Technologie analysiert?

Hierzu finden sich keine Angaben in dem Bericht.

Wurden Aspekte der Generalisierbarkeit/Übertragbarkeit diskutiert?

Auch die Übertragbarkeit wird nicht diskutiert. Für die Übertragbarkeit auf das deutsche Gesundheitssystem könnten die von Leidl et al. [8] oder von Greiner et al. [9] veröffentlichten Kriterien angewendet werden.

Wurden ethische, soziale und juristische Auswirkungen der Technologie berücksichtigt?

Auch hierzu finden sich keine Angaben im Bericht des KCE.

Wurden Angaben zum Peer-Review, zum Revisionsdatum und zu Interessenkonflikten gemacht?

Es finden sich detaillierte Angaben zu den Autoren, hinzugezogenen Experten und Peer-Reviewern sowie möglichen Interessenkonflikten der Beteiligten. Ein Revisionsdatum ist nicht angegeben.

16.2.5 Auflösung des Szenarios

Die Qualität des HTA-Berichts kann auf Grundlage dieser Kriterien insgesamt als hoch eingeschätzt werden, auch wenn nicht alle Aspekte berücksichtigt wurden bzw. Dokumentationsdefizite in den Teilen Wirksamkeit und Sicherheit

bestehen. Die Schlussfolgerungen sind glaubwürdig und nachvollziehbar.

Der HTA-Bericht weist darauf hin, dass auch bei Frauen mit Brustkrebs im Frühstadium der Einsatz von Trastuzumab das ereignisfreie Überleben verbessert, insbesondere durch die Verhinderung des Auftretens von Fernmetastasen (außer Hirnmetastasen). Wird Trastuzumab allerdings nach der Chemotherapie mit Anthrazyklinen eingesetzt, dann besteht ein höheres Risiko einer Herzinsuffizienz. Eine kleinere Studie liefert Hinweise, dass die Wirksamkeit von Trastuzumab vor Einsatz von Anthrazyklinen die Kardiotoxizität reduziert, weshalb die Autoren empfehlen, dieses Therapieregime möglichst umgehend in einer definitiven Studie zu testen. Es gibt keine Evidenz für die Wirksamkeit von Trastuzumab bei Frauen mit Brustkrebs über 70 Jahren. Trastuzumab ist aufgrund der belgischen Datenlage nur für einen Anteil von ca. 10–15% der belgischen Frauen mit HER2-positivem Brustkrebs überhaupt geeignet, und auch nur dann, wenn die LVEF mindestens 55% beträgt. In der ökonomischen Analyse war das Therapieregime Trastuzumab vor der Chemotherapie mit Anthrazyklinen am kosteneffektivsten, allerdings ist hier die Evidenzlage hinsichtlich der Wirksamkeit noch am unsichersten. Der HTA-Bericht empfiehlt, die Therapie mit Trastuzumab zunächst konditional nur im Rahmen qualitätsgesicherter Programme zur Behandlung von Brustkrebs an spezialisierten Zentren unter Beachtung der Ein- und Ausschlusskriterien einzusetzen und ein prospektives Register einzurichten, in dem prognostische Faktoren und Ergebnisse dokumentiert werden.

Aufgrund der Evidenzlage und dem Ergebnis der Kosteneffektivitätsanalyse erarbeiten Sie einen Vertragsentwurf, der sich inhaltlich an die Empfehlung der belgischen Arbeitsgruppe anlehnt. Die Laufzeit wird allerdings zunächst auf ein Jahr befristet, um nach Veröffentlichung noch ausstehender Studienergebnisse eine erneute Bewertung vornehmen zu können.

16.3 HTA in Deutschland

Die Bedeutung von HTA als Grundlage von Entscheidungen im deutschen Gesundheitswesen hat in den letzten Jahren stark zugenommen. Bei Entscheidungen über die Erstattungsfähigkeit von medizinischen Technologien, bei der Ausgestaltung von Disease-Management-Programmen, aber auch im klinischen Alltag können HTA-Berichte hochwertige Informationen zur Verfügung stellen, die zunehmend auch in Anspruch genommen werden. HTA ist als Instrument bei der Gestaltung des Leistungskatalogs im Rahmen der Arbeit des Gemeinsamen Bundesausschusses fest verankert [10]. Grundlage für Vertragsabschlüsse in 2005 zur Sicherstellung der Betreuung von Frauen mit BRCA1/2-Mutationen in Deutschland war ebenfalls ein HTA-Bericht [11]. Im internationalen Vergleich hat HTA in Deutschland einen vergleichsweise hohen Stellenwert.

16.4 Internationale HTA-Aktivitäten

In den 1980er- und frühen 1990er-Jahren breitete sich die Idee von HTA in den westlichen Industrienationen weiter aus: in Nordamerika, Europa, Australien und Neuseeland entstanden öffentlich finanzierte HTA-Institutionen, die sich in unterschiedlich enger Bindung an entscheidungsbefugte Gremien in den jeweiligen Gesundheitssystemen mit der Erstellung von wissenschaftlich fundierter Verfahrensbewertung befassten. Aber auch privatwirtschaftlich (For-Profit) arbeitende und gemeinnützige Institutionen begannen sich mit der Erstellung von HTA-Gutachten zu befassen. Treibende Kräfte dieser Entwicklung waren einerseits die in vielen Gesundheitssystemen knapper werdenden finanziellen Ressourcen bei einem ständig wachsenden Angebot an kostspieligen Technologien, andererseits die insbesondere in-

nerhalb der Disziplin HTA wahrgenommene Verpflichtung, in der Gesundheitsversorgung eine „Evaluationskultur“ zu entwickeln und zu pflegen. Es wurde zunehmend klar, dass eine transparente, streng wissenschaftliche Methodik das Schlüsselement für die Unabhängigkeit, Validität und Glaubwürdigkeit von Verfahrensbewertungen bildet. Diese stellt gleichzeitig eine notwendige Voraussetzung dafür dar, dass Technologiebewertungen zumindest in ihren Kernelementen (der systematischen Sicherheits- und Wirksamkeitsbewertung von Technologien) international übertragbar sind. Vor diesem Hintergrund entstanden eine Reihe von internationalen Kooperationsaktivitäten, die sich einerseits um internationale Zusammenarbeit im Kontext der Erstellung von Verfahrensbewertungen bemühen, andererseits die Implementation von HTA in gesundheitspolitische und versorgungsrelevante Entscheidungsfindungen auch auf internationaler Ebene befördern wollen.

16.4.1 HTA international (www.htai.org)

Die erste internationale Aktivität im Bereich HTA stellte 1985 die Gründung einer internationalen wissenschaftlichen Fachgesellschaft für HTA dar (ISTAHC – International Society for Technology Assessment in Health Care bis 2003; seit 2003 HTAi – HTA international).

HTAi versteht sich als Forum für alle an der wissenschaftlichen und politischen Entwicklung sowie der Anwendung von HTA interessierten Kreise aus Gesundheitsversorgung, Wissenschaft und Wirtschaft. Zu den Mitgliedern gehören sowohl Individuen wie auch Non-for-profit-Organisationen (z.B. öffentlich geförderte HTA-Institute) und For-profit-Organisationen (z.B. Beratungsunternehmen, Hersteller von Pharmazeutika und Medizintechnik). Die Fachgesellschaft sieht ihren Auftrag in der Unterstützung von Entwicklung, Kommunikation, Verständnis und Gebrauch von

HTA weltweit – mit den Zielen, die Einführung von effektiven Technologien in die Versorgung zu fördern und den wirtschaftlichen Einsatz von Ressourcen zu unterstützen.

Die Interest-Subgroups reflektieren die Schwerpunkte der Arbeit innerhalb von HTAi: HTA-Produzenten/Wissenschaft, Industrie, Klinik, Entscheidungsträger, Patienten- und Bürgerbeteiligung, Informationsmanagement, Ethik und HTA für Krankenanstalten. Durch spezielle Förderprogramme (Stipendien für Ausbildungsmaßnahmen, Reisestipendien) unterstützt HTAi insbesondere die Implementation von HTA in Entwicklungs- und Schwellenländern.

Die wichtigsten Kommunikationsinstrumente innerhalb der Fachgesellschaft, aber auch nach außen sind die im Juni stattfindende Jahrestagung und das vierteljährlich bei Cambridge University Press erscheinende *International Journal of Technology Assessment in Health Care (IJTAHC)*. Die Jahrestagungen haben einerseits den Charakter von wissenschaftlichen Tagungen (mit der Gelegenheit, eigene Arbeiten in Vorträgen oder Postern zu präsentieren), andererseits versuchen sie, gezielt die Kommunikation und den Erfahrungsaustausch von HTA-Produzenten, Entscheidungsträgern und Technologieproduzenten herbeizuführen.

Als weitere Kommunikationsplattform dient die öffentlich zugängliche Internetseite der Fachgesellschaft [<http://www.htai.org>] mit weiterführenden Informationen, Kontaktdaten der Mitglieder und Diskussionsforen zu den Arbeitsschwerpunkten.

16.4.2 INAHTA – International Network of Agencies for HTA (www.inahta.org)

Im Gegensatz zu HTAi, die allen Interessen offen steht, handelt es sich beim Netzwerk INAHTA um einen Zusammenschluss von mit öffentlichen Mitteln geförderten

HTA-Institutionen. Es wurde 1993 gegründet und verfügt derzeit (2006) über 45 Mitglieder aus 23 Nationen. Das Sekretariat, das die Koordination des Netzwerks übernimmt und für die Pflege der Webseite verantwortlich ist, ist derzeit beim schwedischen HTA-Institut SBU angesiedelt. INAHTA versteht sich als Forum für die Identifikation und Verfolgung gemeinsamer Interessen öffentlich geförderter HTA-Institutionen. Da der technologische Fortschritt in der Medizin ein globales Phänomen ist, treten in vielen Gesundheitssystemen sehr ähnliche Entscheidungssituationen auf. INAHTAs Arbeit zielt durch Förderung von Kommunikation und Informationsaustausch zwischen den HTA-Institutionen auf die Vermeidung von unnötiger Doppelarbeit. Zu den Informationsstrategien gehören:

- ▲ die Publikationen von englischsprachigen Briefs – strukturierten Kurzfassungen von HTA-Berichten, die von den Mitgliederorganisationen publiziert wurden. Die Kurzfassungen sind auf der INAHTA Internetseite für alle Interessenten einsehbar und enthalten in der Regel einen Internetlink zur Vollversion des Berichtes.
- ▲ die Publikation eines vierteljährlich erscheinenden, englisch-, spanisch- bzw. französischsprachigen Newsletters mit Informationen zu Netzwerkaktivitäten, Aktivitäten der Mitgliederorganisationen und wichtigen Entwicklungen im Bereich HTA allgemein. Auch der Newsletter ist für Nichtmitglieder auf der Webseite einsehbar.
- ▲ Unterhalt und Pflege der HTA-Database in Zusammenarbeit mit dem Center for Reviews and Dissemination des britischen National Health Service. Die kontinuierlich aktualisierte Datenbank enthält strukturierte Abstracts von systematischen Reviews, laufenden und abgeschlossenen Primärstudien mit HTA-Bezug, ökonomischen Evaluationen sowie HTA-Berichten. Alle erfassten Arbeiten stammen aus öffentlich geförderten HTA-Institutionen.

Unter http://www.york.ac.uk/inst/crd/crd_databases.htm kann kostenfrei in der Datenbank recherchiert werden.

Eine weitere Zielsetzung der INAHTA-Arbeit besteht darin, Kernelemente und -aussagen von HTA-Berichten international anwendbar zu machen. Voraussetzung hierfür ist die Anwendung von allseits akzeptierten methodischen Qualitätsstandards. Die INAHTA-Aktivitäten zur Förderung der Qualität von HTA bestehen in der Bereitstellung einer Checkliste zur Qualitätsbewertung von HTA-Berichten (im pdf-Format auf der Webseite erhältlich), der Durchführung von kooperativen HTA-Projekten (bisher: Knochendichtemessung und Osteoporose; Prostatakarzinom-Screening; Telemedizin; Positronenemissionstomographie) und dem Engagement der INAHTA Mitglieder in HTAi und den europäischen HTA-Projekten.

16.4.3 Europäische HTA-Projekte

Seit 1994 wird die Etablierung von HTA als Instrument der Entscheidungsunterstützung in europäischen Gesundheitssystemen durch die europäische Union gefördert.

Von 1994–1997 wurden im Projekt **EURASSESS** Kontakte zwischen HTA-Organisationen der europäischen Länder hergestellt und wichtige Begriffs- und Positionsbestimmungen vorgenommen. Arbeitsziele waren die Harmonisierung von Methoden, Exploration von Disseminationswegen für HTA-Ergebnisse, Auslotung der Möglichkeiten, Kostenübernahme-Entscheidungen an HTA zu koppeln sowie die Entwicklung von Priorisierungsprozessen. Obwohl die Ziele nur teilweise umgesetzt werden konnten, wurde die internationale europäische Kooperation als wichtiges Element für den effizienten Einsatz von HTA auch auf nationaler Ebene identifiziert [12].

Als Folgeprojekt wurde von 1997–1998 das Vorhaben **HTA-Europe** gefördert. Im

Rahmen von HTA-Europe standen vor allem die Förderung der Institutionalisierung von HTA in den europäischen Ländern, der Informationsaustausch und die Entwicklung erster Ideen für eine nachhaltige HTA-Arbeit auf europäischer Ebene auf der Agenda [13].

Mit den Zielen der Förderung einer evidenzbasierten Gesundheitsversorgung in ganz Europa und der Vernetzung von Aktivitäten begann im Jahr 2000 die Förderung des dritten EU-Projekts (European Collaboration for Health Technology Assessment/Assessment of Health Interventions **ECHTA/ECAHI**). Die Projektarbeit wurde von 110 medizinischen und gesundheitspolitischen Experten aus allen (damaligen) 15 EU-Mitgliedsstaaten und acht weiteren beobachtenden Ländern umgesetzt. Sechs Arbeitsgruppen bearbeiteten folgende Themenbereiche: Gesundheitsförderungs-/Präventionsaktivitäten; Etablierung von Routinen im Informationsaustausch; kooperative Assessments; Vorschläge zu methodischem „Best Practice“; Entwicklung von Ausbildungsaktivitäten sowie die Anbindung von HTA-Ergebnissen an gesundheitspolitische Entscheidungsprozesse. Als Hauptergebnis des ECHTA/ECAHI-Projektes wird die Notwendigkeit zur Gründung einer permanenten europäischen Institution oder Kooperationsform formuliert, die die Nachhaltigkeit der Integration von HTA in die nationalen europäischen Gesundheitssysteme unterstützt und gleichzeitig durch Vernetzung die effiziente Arbeit ermöglicht.

Dies umzusetzen ist die übergeordnete Zielsetzung des derzeit (2006–2008) laufenden Projektes European Network for Health Technology Assessment (**EUNetHTA**). An der Projektarbeit sind 59 Organisationen aus 24 EU-Mitgliedsstaaten, zwei Ländern des europäischen Wirtschaftsraums (EEA) sowie die Schweiz beteiligt. Kooperationsbeziehungen bestehen zu HTA-Organisationen aus den USA, Kanada, Australien und Israel, zur Weltgesundheitsorganisation, der OECD, dem europäischen Rat, INAHTA, HTAi, EuroScan,

zum internationalen Leitliniennetzwerk GIN und zur Cochrane Collaboration. Die Projektleitung wird von der dänischen HTA-Agentur DACEHTA wahrgenommen [14].

In den drei Jahren der Projektlaufzeit sollen

- ▲ die Möglichkeiten für mehrdimensionalen und belastbaren Input in gesundheitsrelevante Entscheidungsfindungen etabliert werden,
- ▲ Doppelarbeiten innerhalb Europas reduziert und dadurch Ressourcen im Bereich HTA eingespart werden,
- ▲ ein besseres Verständnis für die Zusammenhänge zwischen HTA und Entscheidungsfindung in den Mitgliedsstaaten entwickelt und
- ▲ Länder mit begrenzten Erfahrungen mit HTA unterstützt werden.

Die Arbeit in sechs Arbeitspaketen (neben der Projektleitung und Evaluation) umfasst die Themen Kommunikation, Methodik, Adaptation von internationalen Ergebnissen auf nationale Gegebenheiten, Transfer von HTA-Ergebnissen in Entscheidungsprozesse, Identifikation/Monitoring von innovativen Technologien sowie die Unterstützung von Ländern mit unzureichender HTA-Infrastruktur. Aktuelle Informationen können auf der Internetseite des Projekts unter www.eunetha.eu eingesehen werden.

Literatur

- [1] Henshall C et al., Priority setting for health technology assessment. Theoretical considerations and practical approaches. Priority setting Subgroup of the EUR-ASSESS Project. *Int J Technol Assess Health Care.* (1997), 13, 144–185
- [2] European Medicines Agency (EMA). European public assessment report Herceptin. <http://www.emea.europa.eu/humandocs/Humans/EPAR/herceptin/herceptinM.htm> (24.11.06)
- [3] Huybrechts M, Hulstaert F, Neyt M, Vrijens F, Ramaekers D. Trastuzumab in early stage breast cancer. *Health Technology Assessment (HTA).* 2nd ed. http://kce.fgov.be/index_en.aspx?ID=0&SGREF=5211&CREF=7191 (24.11.06)
- [4] Raftery J, NICE: faster access to modern treatments? *Analysis of guidance on health technologies.* *BMJ.* (2001), 323, 1300–1303
- [5] Stevens A, Milne R, Health technology assessment in England and Wales. *Int J Technol Assess Health Care.* (2004), 20, 11–24
- [6] Busse R et al., Best practice in undertaking and reporting health technology assessments. *Int J Technol Assess Health Care.* (2002), 18, 361–422
- [7] Perleth M (2003) Evidenzbasierte Entscheidungsunterstützung im Gesundheitswesen. Konzepte und Methoden der systematischen Bewertung medizinischer Technologien (Health Technology Assessment) in Deutschland. WiKu-Verlag, Berlin
- [8] Leidl R, von der Schulenburg JM, Wasem J (1999) Ansätze und Methoden der ökonomischen Evaluation-eine internationale Perspektive. Nomos, Baden-Baden
- [9] Greiner W, Schöffski O, von der Schulenburg JM, The transferability of international economic health-economic results to national questions. *Health Econ Prev Care.* (2000), 1, 94–102
- [10] Niebuhr D et al. (2004) Die Bestimmung des Leistungskataloges in der gesetzlichen Krankenversicherung. Band 2: Verfahren und Kriterien zur Bestimmung des Leistungskatalogs in der gesetzlichen Krankenversicherung vor dem Hintergrund internationaler Erfahrungen. Hans-Böckler-Stiftung, Düsseldorf
- [11] Gerhardus A et al. (2005) BRCA-Erblicher Brust- und Eierstockkrebs. Beratung-Testverfahren-Kosten. Springer, Berlin
- [12] Sassi F, The European way to health technology assessment. Lessons from an evaluation of EUR-ASSESS. *Int J Technol Assess Health Care.* (2000), 16, 282–290
- [13] Jonsson E et al., Executive summary of the ECHTA/ECAHI project. European Collaboration for Health Technology Assessment/Assessment of Health Interventions. *Int J Technol Assess Health Care.* (2002), 18, 213–217
- [14] Kristensen FB, Chamova J, Hansen NW, Toward a sustainable European Network for Health Technology Assessment. The EUnetha project. *Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz.* (2006), 49, 283–285

17 Kritische Bewertung von gesundheitsökonomischen Studien

Reiner Leidl

Kernaussagen

- ▲ Die Medizin kann ihren Handlungsspielraum durch die Beachtung der Wirtschaftlichkeit vergrößern.
- ▲ International gibt es zahlreiche Studien zur Wirtschaftlichkeit von medizinischen Leistungen, und es liegen anerkannte Qualitätsstandards für die ökonomische Evaluation vor.
- ▲ Die methodische Brauchbarkeit der Instrumente für die Messung von Kosten und von gesundheitlichen Ergebnissen ist zu prüfen.
- ▲ Bei der Bewertung der ökonomischen Evidenz spielt der Entscheidungskontext, in dem die Evidenz genutzt werden soll, eine wichtige Rolle.

17.1 Ein Szenario

In einem Fachblatt lesen die Chefin einer Universitätsklinik und der Vorstand einer großen Krankenkasse von der Entwicklung eines neuen DNA-Chips, mit dem diejenigen zehn Prozent von Patienten einer Krebsart, bei denen eine schon länger verfügbare, teure Chemotherapie die Neubildung verschwinden lässt, identifiziert werden können. Der Hersteller der Therapie verweist auf seine Forschungskosten und eine neue amerikanische Studie, die – ohne Chip – die Kosten-Effektivität seines Präparats gezeigt hatte, und kündigt eine deutliche Preiserhöhung an. Die Ärztin überlegt, wie die Hinweise zur Wirtschaftlichkeit des Behandlungsverfahrens in den Leitlinien ihrer Fachgesellschaft verändert werden müssen, und der Kassenvorstand fragt sich, ob bei den neuen Therapiepreisen das Wirtschaftlichkeitsgebot noch erfüllt ist.

17.2 Warum die Wirtschaftlichkeit medizinischer Leistungen untersuchen?

Die Wirtschaftlichkeit medizinischer Leistungen kann mit Methoden der ökonomischen Evaluation untersucht werden, die ein Teilgebiet der Gesundheitsökonomie darstellt. Die Informationen zur Wirtschaftlichkeit der medizinischen Leistungen sollen helfen, wirtschaftliche Entscheidungen in der Gesundheitsversorgung rational treffen zu können und die Effizienz der Versorgung zu verbessern. Wie andere Lebensbereiche unterliegt auch medizinisches Handeln dem Phänomen der Knappheit: Die verfügbaren Ressourcen reichen nicht aus, um alles Mögliche zu realisieren. Den bestmöglichen Umgang mit der Knappheit bezeichnet man in der Ökonomie als Effizienz. Um sie zu erreichen, müssen – vereinfacht – zwei Bedingungen erfüllt werden:

- ▲ Technische Effizienz. Sie wird dann erreicht, wenn das Wirtschaftlichkeitsprinzip erfüllt ist, d.h. wenn mit einem gegebenen Input an Ressourcen ein Maximum an Output – etwa die Verbesserung der Gesundheit – erreicht wird oder wenn ein gegebenes Outputziel mit einem minimalen Ressourceneinsatz erzeugt wird.
- ▲ Allokative Effizienz. Für sie wird zusätzlich gefordert, dass diejenigen Güter und Dienstleistungen – etwa Gesundheitsleistungen – erstellt werden, die von gut informierten Patienten und Versicherten gewünscht werden und für welche diese, zum Beispiel über ihre Versicherungsprämien, zu zahlen bereit sind.

Um die Effizienz von Gesundheitsleistungen zu beurteilen, sind daher deren Kosten und ihr gesundheitliches Ergebnis einzuschätzen. Ferner ist zu prüfen, ob seitens eines Leistungsnachfragers die Bereitschaft besteht, die Kosten für die gesundheitliche Verbesserung zu tragen. Bei knappen Budgets können Entscheidungsträger auf der Basis solcher Informationen die Effizienz der Versorgung verbessern. Die Medizin kann ihren künftigen Handlungsspielraum durch die Beachtung der Wirtschaftlichkeit vergrößern.

17.3 Wo setzt die ökonomische Evaluation an?

Generell kann die ökonomische Evaluation gut am Goldstandard der Untersuchung der medizinischen Effektivität anknüpfen, nämlich am randomisierten klinischen Versuch. Zugleich muss sie – anders als die Analyse der medizinischen Effektivität – spezifische Elemente desjenigen Gesundheitssystems abbilden, dem sie Informationen für ökonomische Entscheidungen liefern möchte. Das Beispiel zur Wirtschaftlichkeit eines Arzneimittels, das Krankenhauseinweisungen reduziert, verdeutlicht dies: So muss bei gleichem Medikamentenpreis in Deutschland das Mehrfache an Krankenhaustagen eingespart werden wie in den USA, um unter sonst gleichen Bedingungen die gleiche Wirtschaftlichkeit zu erreichen, da der amerikanische Krankenhaustag um ein Vielfaches teurer ist.

Neben der systemspezifischen Ausrichtung versucht man in der ökonomischen Evaluation, möglichst alle langfristigen Auswirkungen auf Kosten und Effekte einzubeziehen. Klinische Versuche können häufig diese Zeiträume nicht abdecken. In der ökonomischen Evaluation wird daher vermehrt die Modellierung eingesetzt. Als Alternative zum Experiment und zur Beobachtung stellen solche Modelle quantitative Entwicklungen und Zusammenhänge dar (wie z.B. die

Übergangswahrscheinlichkeiten zwischen Krankheitsstadien in Abhängigkeit von der Therapie und die damit verbundenen Kosten). Die Daten eines Modells können auf eigenen oder fremden Messungen oder auf Annahmen beruhen; dabei sollte in nachvollziehbarer Weise die beste verfügbare Evidenz einbezogen werden.

17.4 Welche Standards gibt es für ökonomische Evaluationen?

Im Mai 2006 dokumentierte die Health Economic Evaluations Database des Londoner Office of Health Economics (www.ohe-heed.com) weltweit 33 000 Arbeiten zur ökonomischen Evaluation in allen Bereichen des Morbiditätsspektrums, darunter etwa die Hälfte empirische Studien. Wie ein Blick in die Literaturdatenbank PubMed zeigt, werden etwa in den USA, in Großbritannien und in den Niederlanden deutlich mehr ökonomische Evaluationsstudien durchgeführt als in Deutschland. In diesen Ländern bestehen seit Jahren weit entwickelte Forschungsstrukturen auf dem Gebiet der Wirtschaftlichkeitsuntersuchung [1].

Die Verbreitung der Wirtschaftlichkeitsstudien begann erst im letzten Viertel des vergangenen Jahrhunderts, also zwei bis drei Jahrzehnte nach der Diffusion der klinischen Studien, deren Qualitätskriterien entsprechend länger ausgereift sind. International sind heute auch die methodischen Grundlagen der ökonomischen Evaluation weitgehend akzeptiert. Zur ökonomischen Evaluation liegen zahlreiche Lehrbuchbeiträge vor (z.B. [2], in deutscher Sprache [3] oder [4]). Zur Durchführung von Studien gibt es aufgrund der systemspezifischen Erfordernisse verschiedene Richtlinien (z.B. für Einzelstudien die von einem Lehrbuch begleitete amerikanische Version von [5] und die Vorgaben des British Medical Journal [6], speziell für Modellierungsstudien [7] und für Studien-

übersichten die Qualitätskriterien einer Konsensusgruppe gemäß [8]). Eine deutsche Konsensusgruppe hat 1999 eine zweite Version der relativ allgemeinen deutschen Richtlinien verabschiedet (in [1]), ein aktuellerer Standard stammt von einer Arbeitsgruppe aus der Rehabilitationsforschung [9]. Internationale Erfahrungen, etwa aus Australien, wo die ökonomische Bewertung von Arzneimitteln bereits in den 90er-Jahren routinemäßig etabliert wurde, zeigen, dass Studien einige Zeit nach der Einführung von Richtlinien noch häufig dagegen verstießen [10].

Anders als bei der medizinischen Evidenz können ökonomische Evaluationsstudien nicht unabhängig von ihrer Nutzung präzise bewertet werden: Der Entscheidungskontext, in den die ökonomische Evidenz einfließen soll, bestimmt mit, ob die verwendeten Methoden angemessen sind. Leistungsfähige Standards für die ökonomische Evaluation müssen daher methodisch genau differenziert und auf den relevanten Entscheidungskontext ausgerichtet sein. Anstel-

le eines bestimmten Standards wird hier ein genereller Überblick über die zentralen Komponenten einer ökonomischen Evaluationsstudie gegeben: Tabelle 17.1 benennt wesentliche Kriterien, die in einem qualitativ hochwertigen Beitrag transparent dargestellt sein sollten.

17.5 Welche Arten der ökonomischen Evaluation gibt es?

Sieht man vom Typus der Krankheitskostenstudie ab, die sich mit den gesamten Kosten eines Krankheitsbilds befasst, gibt es in der ökonomischen Evaluation vier Studientypen. Alle befassen sich mit der Wirtschaftlichkeit von Interventionen, unterschieden sich aber nach der Art, in der die Effektseite gemessen wird (s. Abb. 17.1). In jedem Fall wird von einem Vergleich ausgegangen, bei dem eine Interventionsgruppe einer Gruppe ohne Intervention oder mit herkömmlicher Behandlung gegenübergestellt wird. Im Fall der Kosten-Mi-

Tab. 17.1: Zentrale Komponenten einer ökonomischen Evaluationsstudie

Studienteil	Einzelkomponenten
Fragestellung und Evaluationsrahmen	Beschreibung von Untersuchungsziel, Zielpopulation, zu vergleichende Intervention und Alternativen, Evaluationsart und Perspektive
Untersuchungsansatz	Beschreibung des experimentellen Designs, der Beobachtungsstudie oder des Modellierungsansatzes; der Messkonzepte für Kosten und Effekte; Bezugsjahr und Zeithorizont
Datenerfassung	Beschreibung von Art, Quelle und Erfassungsmethode der Daten: (i) auf der Kostenseite der Ermittlung von Mengen und Preisen (ii) auf der Effektseite der Beschreibung und Bewertung von Gesundheitszuständen bzw. anderer Effektgrößen (iii) der Diskontierung von Kosten und Effekten
Analyse und Ergebnisse	Darstellung der Basisanalyse: absolute und inkrementelle Kosten und Effekte sowie deren Relation, Analyse der Gesamtpopulation und wichtiger Subgruppen; Behandlung von Unsicherheiten: uni- und multivariate Sensitivitätsanalyse oder stochastische Analyse
Diskussion und Schlussfolgerungen	Diskussion von Ansatz, Umsetzung, Datenqualität, Ergebnissen und Einschränkungen der Studie, Interpretation und Generalisierbarkeit der Evidenz, verteilungspolitische Aspekte; Ansatzpunkte für Entscheidungen

Anmerkung: Ferner sollte die Interessenposition der Autoren ausgewiesen werden.

nimierungs-Analyse wird nur die Kostenseite betrachtet, was impliziert, dass von einem Erreichen identischer Effektniveaus in beiden Vergleichsgruppen ausgegangen wird. Die häufigste Evaluationsart ist die Kosten-Effektivitäts-Analyse, bei der auf der Effektseite ein zentraler klinischer Parameter zum Vergleich herangezogen wird. Dies könnte z.B. die Lebensdauer, das Auftreten eines schwerwiegenden kardialen oder cerebralen Ereignisses oder die Lipidkonzentration im Blut sein. Nachteil dieser Evaluationsart ist, dass klinische Parameter oft nicht alle für den Patienten wesentlichen Aspekte des gesundheitlichen Ergebnisses erfassen. Bei der Kosten-Nutzwert-Analyse werden alle Veränderungen der gesundheitsbezogenen Lebensqualität über die Beobachtungszeit in einem als Nutzwert charakterisierten Indikator gemessen. Bei methodisch korrekter Durchführung kann mit diesem Studientypus besonders hochwertige Evidenz erzeugt werden; seine Bedeutung hat in den letzten Jahren zugenommen.

Am häufigsten mit der ökonomischen Evaluation in Verbindung gebracht wird die Kosten-Nutzen-Analyse. Dies entspricht jedoch nicht der international akzeptierten Fachterminologie. Diese fordert, dass in Kosten-Nutzen-Analysen alle gesundheitlichen Effekte in Geldeinheiten bewertet werden.

Für solche monetären Bewertungen liegen zahlreiche Studienbeispiele und Methodenarbeiten vor, doch aufgrund der hohen methodischen Anforderungen, die mit einer brauchbaren und fundierten Bewertung der Gesundheit in Geldeinheiten verbunden sind, findet man diesen Studientypus empirisch am seltensten vor.

Bei allen vier Studientypen sollte der Vergleich zwischen Interventions- und Kontrollgruppe so durchgeführt werden, dass den zusätzlich durch die Intervention verursachten Kosten die zusätzlich erreichten Effekte gegenübergestellt werden; man spricht dann auch von einer inkrementellen Analyse.

17.6 Wie werden die Kosten gemessen?

Jede Kostenmessung geht in drei Schritten vor: Identifikation der zu messenden Kategorien, Messung der verbrauchten Ressourceneinheiten und monetäre Bewertung je verbrauchter Einheit. Bei den Kosten einer neuen Technologie mag eine Krankenversicherung primär an ihre Ausgaben denken, ein Patient an seine Selbstzahlung und ein Krankenhausarzt, der unter Fallpauschalen wirtschaftet, an die Betriebskosten seiner Ab-

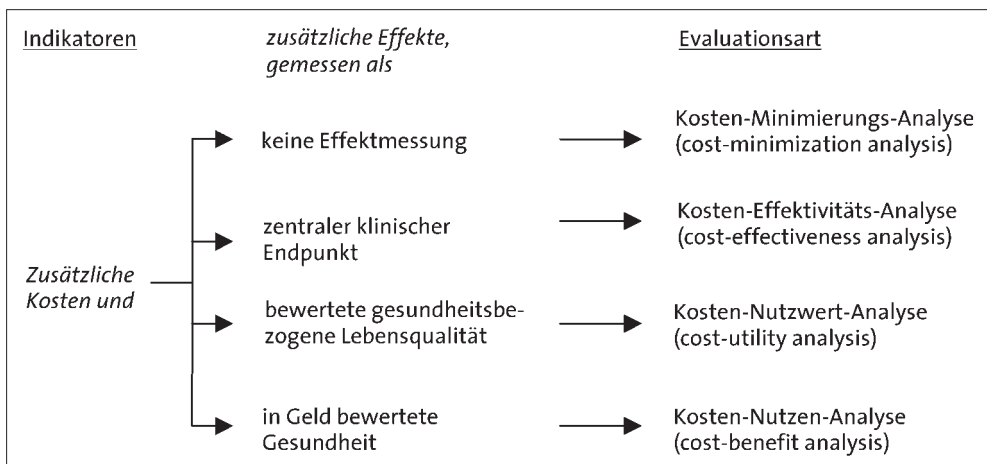


Abb. 17.1: Indikatoren und Arten der ökonomischen Evaluation

teilung. Die Antwort auf die Frage, welche Kosten einzubeziehen sind, hängt von der Perspektive der Evaluation ab. Die beiden wichtigsten sind

- ▲ die gesamtgesellschaftliche Sicht, aus der jeglicher Ressourcenverbrauch einzubeziehen und zu den für die Gesellschaft relevanten Preisen zu bewerten ist und
- ▲ die Perspektive des Finanzierungsträgers, etwa der Krankenkassen, aus der nur die Kostenkategorien einzubeziehen sind, für die der Finanzierungsträger auch aufzukommen hat; sie sind in entsprechender Höhe zu bewerten.

Krankheitsbedingte Pflege durch Angehörige stellt beispielsweise aus gesamtgesellschaftlicher Sicht einen zu bewertenden Ressourcenverbrauch dar. Aus Sicht einer Krankenkasse, die hierfür nicht aufzukommen hat, wäre diese Größe nicht in die Untersuchung mit einzubeziehen. Viele Ökonomen befürworten eine Analyse aus gesamtgesellschaftlicher Sicht, da diese keine Partikularinteressen widerspiegelt.

Kostenanalysen können eine Vielzahl unterschiedlicher Kostenarten einbeziehen. Grob unterschieden werden direkte Kosten der medizinischen Versorgung, die sich etwa auf die ambulante und stationäre Betreuung oder den Arzneimittelverbrauch beziehen können, direkte nicht-medizinische Kosten, die sich etwa auf häusliche Unterstützung beziehen können, und indirekte Kosten. Die letztgenannte Kategorie beschreibt im Wesentlichen den krankheitsbedingten Ausfall an wirtschaftlicher Produktion und wird meist durch entgangene Arbeitszeit und durchschnittlichen Arbeitslohn bewertet. Ein Verzicht auf die Messung von Unterschieden im krankheitsbedingtem Produktionsausfall kann ein wichtiges Ziel des medizinischen Handelns übersehen, etwa, wenn durch medizinische Rehabilitation die Gesundheit und Erwerbsfähigkeit wieder hergestellt werden sollen. Da unterschiedliche

Gruppen in der Gesellschaft unterschiedlich stark am Arbeitsmarkt beteiligt sind und unterschiedliche durchschnittliche Arbeitslöhne erzielen, kann ein Vergleich der Kosten von Interventionen, die unterschiedliche Zielgruppen betreffen, diskriminierend wirken. Bei der Entscheidung über die Art und Weise des Einbezugs von indirekten Kosten kommt es daher auf die Untersuchungsfragestellung an.

Bei der Bewertung des Ressourcenverbrauchs wird in der Praxis vielfach mit Gebührenordnungen, Fallpauschalen und teilweise auch allgemeinen Durchschnittswerten gearbeitet. Welcher Differenzierungsgrad und welche Bewertungsform die „richtige“ ist, kann nicht generell festgelegt werden. Vielmehr ist zu prüfen, ob das Kostenmesskonzept der formulierten Fragestellung und den mit den Ergebnissen zu unterstützenden Entscheidungen adäquat ist. Die pauschale Bewertung des Ereignisses „Durchführung einer interventionellen kardiologischen Maßnahme“ im Rahmen eines Studienvergleichs zur langfristigen Auswirkung des Einsatzes von Vasodilatoren kann angemessen sein, ein Vergleich des zugehörigen Einsatzes verschiedener Stents würde jeweils eine differenzierte Kostenermittlung erfordern.

Generell können die einbezogenen Kostenarten, ihre Differenzierung und die eingesetzten Mess- und Bewertungsverfahren die Ergebnisse einer Kostenmessung beeinflussen. Kostendaten können aus Routinestatistiken, aus dem Rechnungswesen von Krankenhäusern und Krankenversicherungen oder über eigene Leistungserhebungen erschlossen werden. Bei Beteiligung verschiedener Versorgungsinstitutionen und von Patientengruppen, die unterschiedlichen Versicherungen angehören, können Befragungen über sogenannte Kostenwochenbücher eine schlüssige Messung ermöglichen [11]. In jedem Fall muss sich eine Kostenmessung den Fragen der Methodenqualität, insbesondere der Validität, Reliabilität und Reagibilität stellen [3].

Eine nachvollziehbare Wirtschaftlichkeitsanalyse muss bei der Kostenmessung genau angeben, welches Verfahren eingesetzt wurde und wie die einzelnen Komponenten erhoben, bewertet und aggregiert wurden. Dann kann geprüft werden, ob die Studienmethodik den Qualitätsnormen entspricht und ob die gewählte Methodik für den Entscheidungszusammenhang brauchbar ist.

Bei einer Verrechnung von Kostengrößen über die Zeit ist zu berücksichtigen, dass ein fester Betrag zu unterschiedlichen Zeitpunkten – etwa heute sowie in fünf Jahren – nicht denselben Wert besitzt. Grund hierfür ist die Möglichkeit, den Betrag in der Zwischenzeit zu nutzen, etwa um Zinsen zu erzielen. Eine Verrechnung der Kosten zu unterschiedlichen Zeitpunkten kann über eine Diskontierung erfolgen. Sie nimmt mithilfe eines Zinssatzes, der Diskontrate, eine zeitliche Angleichung der Beträge vor. Insbesondere bei Evaluationen über einen langen Zeitraum, etwa bei Therapie chronischer Erkrankungen, können unterschiedliche Diskontraten zu unterschiedlichen Ergebnissen führen. Die verwendete Rate sollte daher begründet und Alternativen dargestellt werden. Will man das Verhältnis von Kosten und Effekten zu unterschiedlichen Zeitpunkten nicht durch eine alleinige Diskontierung der Kosten verzerren, ist es erforderlich, auch eine Diskontierung der Effektseite vorzunehmen.

17.7 Wie wird das gesundheitliche Ergebnis gemessen?

Hier wird nur die wichtigste zusammenfassende Effektgröße betrachtet, die gesundheitsbezogene Lebensqualität. Um die gesundheitlichen Effekte einer medizinischen Maßnahme messen zu können, müssen Veränderungen in den verschiedenen Dimensionen dieser Lebensqualität und im Gesamtzustand erfasst werden. Um die gesundheitliche Entwicklung über die Zeit hinweg beschreiben können, muss auch noch die Zeit berücksichtigt werden, über welche die jeweiligen Gesundheitszustände anhalten (s. Tab. 17.2). Der resultierende Indikator beschreibt dann den Nutzwert.

Zu den wesentlichen Dimensionen der gesundheitsbezogenen Lebensqualität gehören das physische, psychische und soziale Funktionieren bzw. die Rollenerfüllung und dabei mögliche Einschränkungen. Beispielsweise erfasst der EuroQol 5D (zur Forschungsgruppe und Übersicht über die zahlreichen Anwendungen s. www.euroqol.org), das einzige derzeit zu Kosten-Nutzwert-Studien in Deutschland eingesetzte Instrument, in seinem beschreibenden Teil die Dimensionen Beweglichkeit, Selbstversorgung, allgemeine Tätigkeiten, Schmerz und Angst/Depression. Als weitere Instrumente werden z.B. der kanadische Health Utility Index ein-

Tab 17.2: Messkomponenten des QALY-Ansatzes

	Beschreibung	Bewertung	Zeitkomponente	Resultat
Komponente	Angaben zu Items der gesundheitsbezogenen Lebensqualität	Bewertung des Gesundheitszustands mit einem Wert zwischen 0 und 1 (für vollständige Gesundheit)	Multiplikation der Bewertung mit der Dauer des Gesundheitszustands	Qualitätsgleiche Lebensjahre (QALYs)
Beispiele	EuroQol 5D, Health Utility Index, SF-12	Visuelle Analogskala, Standardlotterie, Methode der zeitlichen Abwägung	Bewertung 0,25 mal Dauer von 4 Jahren	1 QALY
Ausführende	Patienten	Patienten oder Allgemeinbevölkerung	Studienpersonal	Patienten- oder Populationsrepräsentativität

gesetzt, der auch die Funktionsfähigkeit der Sinne einbezieht, oder der in der Medizin häufig verwendete SF-36 beziehungsweise seine Kurzformen (vgl. [3]). Im Allgemeinen können die Beschreibungen zur gesundheitsbezogenen Lebensqualität durch die Studienpatienten selbst ausgefüllt werden, denkbar wäre in Fällen, in denen dies nicht möglich ist, auch ein Ausfüllen durch medizinisches Fachpersonal.

Den nächsten Schritt bildet die Bewertung des beschriebenen Gesundheitszustands in einem Index. Es gibt drei sehr häufig gebrauchte Verfahren. Die visuelle Analogskala gleicht einem Thermometer, das vom schlechtesten denkbaren Gesundheitszustand beim Indexwert 0 bis zum besten denkbaren Gesundheitszustand beim Indexwert 100 reicht und auf dem der Proband direkt seine Wertung einträgt. Das Verfahren der Standardlotterie stellt in der einfachsten Form den Probanden vor die Alternative, in seinem gegenwärtigen Gesundheitszustand zu verbleiben oder sich einem Eingriff zu unterziehen, der mit einer Wahrscheinlichkeit (p) zur Wiederherstellung der vollständigen Gesundheit und mit der Gegenwahrscheinlichkeit ($1-p$) zum sofortigen Tod führt. Durch Variation dieser Wahrscheinlichkeiten wird der Punkt ermittelt, an dem der Proband zwischen beiden Alternativen indifferent ist; die Auftretenswahrscheinlichkeit der vollständigen Gesundung gibt dann zugleich die Bewertung des gegenwärtigen Gesundheitszustands des Patienten wieder. Bei der Methode der zeitlichen Abwägung wägt der Proband ab, im gegenwärtigen Gesundheitszustand eine vorgegebene Zeit (t) zu verbleiben, oder einen kürzeren Zeitraum (t^*) im Zustand vollständiger Gesundheit leben zu können. An dem Punkt, den dem der Proband indifferent zwischen beiden Zeiträumen ist, beschreibt der Indikator $(t^*)/(t)$ ebenfalls eine Einstufung des Gesundheitszustands auf einer Skala von 0 bis 1.

Zu diesen Bewertungsverfahren finden sich in der Literatur ausführliche theoretische

und methodische Untersuchungen. Die drei Verfahren eignen sich nicht gleichermaßen für jede Bewertungssituation. Da die visuelle Analogskala – im Unterschied zu den beiden anderen Verfahren – keine Alternative vorsieht, und beim Verfahren der Standardlotterie aufgrund der vorgegebenen Unsicherheit zusätzlich noch der Grad der Risikoaversion des Patienten eine Rolle spielt, messen die Verfahren unterschiedliche Konstrukte der Bewertung von Gesundheit. Empirisch führt dies zu unterschiedlichen Wertungen des gleichen Gesundheitszustands. Direkt verglichen werden können nur Studien mit den gleichen Meßmethoden.

Unterschiede in der Wertung können auch durch die Person entstehen, welche die Wertung durchführt. Anstelle Wertungen von Patienten heranzuziehen, fordern Richtlinien in einigen Ländern, die Bewertung durch die Allgemeinbevölkerung vornehmen zu lassen. Begründet wird dies damit, dass die Ergebnisse der wirtschaftlichen Beurteilung zur Entscheidung über die Mittel aus kollektiven Budgets der sozialen Sicherung verwendet werden. Generell sollte die Effektmessung in hochwertigen Studien genau beschrieben und begründet sein. Hierzu gehört auch eine Prüfung der Eignung der eingesetzten Messverfahren bei der jeweiligen Untersuchungsgruppe. Nur durch Vergleiche mit klinisch akzeptierten Effektmaßen und eine Reihe psychometrischer Tests kann geprüft werden, ob die Messverfahren bei der Zielgruppe valide, reliabel und sensitiv eingesetzt werden können, und oft fehlen diese Prüfungen.

Die Wertung des Gesundheitszustands muss dann mit seiner Dauer kombiniert werden. Der bekannteste Ansatz verwendet eine multiplikative Verknüpfung von Wertung und Zeit. Durch die Qualitätsgewichtung der Zeit entsteht die Rechengröße ‚qualitätsgleiche Lebensjahre‘ (Quality-adjusted life years, QALYs). Ein QALY stellt ein bei vollständiger Gesundheit verbrachtes Lebensjahr dar. Auch für die Berechnung der QALYs ist die

Erfüllung einer Reihe von Annahmen erforderlich, auf die hier nicht weiter eingegangen wird. Obwohl Ärzte bei ihren Therapieentscheidungen letztlich immer eine Wertung der gesamten gesundheitlichen Effekte vornehmen, stehen sie den Schritten der QALY-Messung recht kritisch gegenüber [12].

17.8 Was ist das Ergebnis der ökonomischen Evaluation?

Schließlich müssen die zusätzlichen Kosten und Effekte in der zentralen Zielgröße, der Kosten-Effektivitäts-Relation zusammengebracht werden (s. Abb. 17.2). Sind die zusätzlichen Effekte einer neuen Intervention positiv und gehen die Kosten sogar zurück, wäre der Idealfall erreicht: die neue Intervention wäre dominant und nach dem Wirtschaftlichkeitsprinzip immer durchzuführen. Steigen die Kosten und kommt es zu negativen zusätzlichen Effekten, würde sich niemand für diese Intervention entscheiden wollen. In den meisten Fällen sind jedoch zusätzli-

che positive Effekte auch mit zusätzlichen Kosten verbunden. Dann ist eine zusätzliche Wertung erforderlich, ob die für die Effektsteigerung notwendigen Kosten auch getragen werden sollen. Im System der sozialen Sicherung könnte diese Wertung beispielsweise von der Versicherungsgemeinschaft oder vom zuständigen Ministerium vorgenommen werden; in anders organisierten Gesundheitssystemen könnte die Verantwortung für die Wertung bei anderen Entscheidungsträgern liegen. Für die tatsächliche Entscheidung werden dabei auch andere Aspekte als die Wirtschaftlichkeit, etwa die Pflicht zur Hilfeleistung als ethische Kategorie, eine Rolle spielen. Es ist nicht Aufgabe der Gesundheitsökonomie, diese Wertung vorzunehmen; sie stellt lediglich die Informationen zur Wirtschaftlichkeit zur Verfügung. Zur Entscheidungsunterstützung kann sie zusätzlich noch auf die Wirtschaftlichkeit von Interventionen hinweisen, in denen bereits durch die Entscheidungsträger Wertungen vorgenommen wurden.

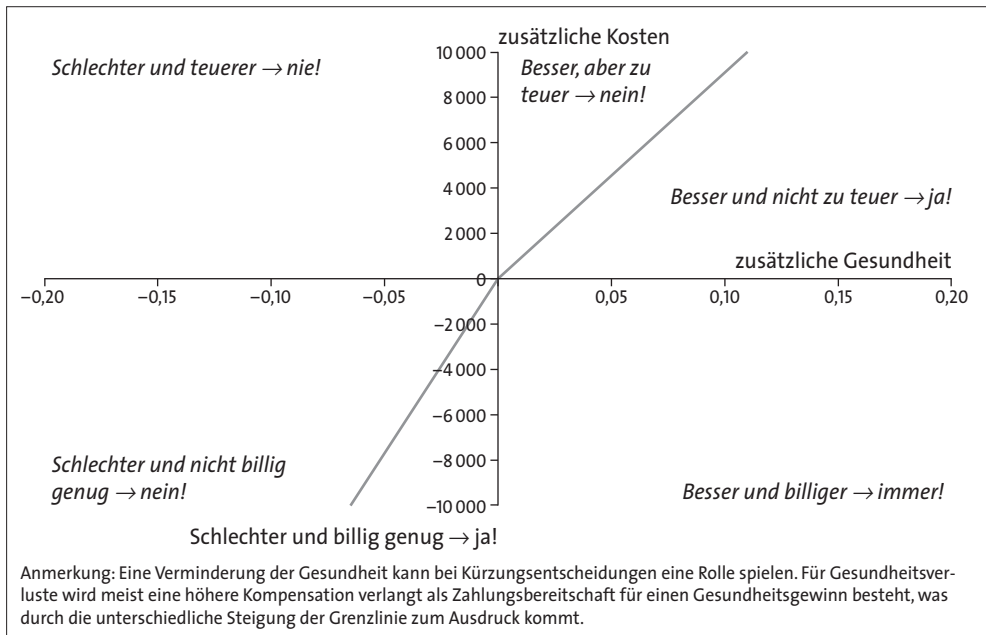


Abb. 17.2: Rationale Entscheidung in der Kosten-Effektivitätsebene

17.9 Kann Unsicherheit berücksichtigt werden?

Vor einer abschließenden Interpretation der Studienergebnisse sind die Unsicherheiten, die dem Ergebnis zugrunde liegen können, zu analysieren. In Fällen, in denen in einer experimentellen Studie prospektiv Daten der Kosten- und der Effektseite für alle Patienten gesammelt wurden, kann die Kosten-Effektivitäts-Relation einer stochastischen Analyse unterzogen werden. Da es sich bei der Zielgröße um eine Relation handelt, resultieren hierbei Konfidenzellipsen. Für die Berechnung der Konfidenzintervalle wird eine Reihe von Ansätzen diskutiert; darüber hinaus kann man die Streuung der Untersuchungsergebnisse in Kosten-Effektivitäts-Akzeptanzkurven darstellen, die zeigen, mit welcher Wahrscheinlichkeit eine selbstgewählte Kosten-Effektivitäts-Grenze nicht überschritten wird [13].

Beruhend auf den Untersuchungsergebnissen hingegen auf Annahmen oder auf unzureichend genauen bzw. repräsentativen Messungen, müssen Sensitivitätsanalysen durchgeführt werden. In der univariaten Analyse sind dabei Parameterannahmen innerhalb begründeter Intervalle zu variieren und die Auswirkungen auf die Zielgröße und ihre Bestandteile zu beschreiben. Bei multivariaten Sensitivitätsanalysen werden mehrere Parameterannahmen gleichzeitig variiert, wobei die Auswahl und Kombination der Variablen zu begründen ist, etwa mit den Resultaten der univariaten Analyse. Der rechenaufwendige Ansatz der Monte-Carlo-Simulation variiert alle relevanten Parameter gleichzeitig, wobei diese aus ihren jeweils vorgegebenen Verteilungen zufällig ausgewählt werden. Ergebnisse von Evaluationsstudien, die sich im Rahmen der begründeten Sensitivitätsanalyse qualitativ nicht verändern, können als stabil angesehen werden. In anderen Fällen sind die wesentlichen Einflussfaktoren auszuweisen und deren mögliche Relevanz für Entscheidungen zu prüfen.

17.10 Was ist ökonomische Evidenz?

Den wissenschaftlichen Kenntnisstand zur Wirtschaftlichkeit einer medizinischen Leistung kann man als ökonomische Evidenz bezeichnen. Wenn mehrere Studien verfügbar sind, lassen sie sich miteinander dann gut vergleichen, wenn sie mit methodisch hoher Qualität durchgeführt und transparent dokumentiert wurden; dies kann auch sehr differenziert untersucht werden [1]. Die Brauchbarkeit der Evidenz hängt auch von ihrem Nutzungszweck ab. So sollen die Resultate ökonomischer Evaluationsstudien in Entscheidungen im Gesundheitssystem eingehen. Wegen der Unterschiede zwischen den Systemen – etwa in der Zuständigkeit für die Versorgung und in der Höhe der Gesundheitsausgaben – ist es nicht immer möglich, die ökonomische Evidenz aus einem Land als Informationsbasis in einem anderen Land zu verwenden. Beispielsweise könnte eine neue Technologie, die das gleiche Gesundheitsergebnis wie eine herkömmliche Technologie erzielt, in einem Land Kosten einsparen; sie wäre dort immer durchzuführen. Würde sie hingegen in einem anderen Land mit anderem Preisgefüge zu Kostensteigerungen führen, wäre sie dort grundsätzlich abzulehnen. Anhand einer Reihe von Kriterien kann geprüft werden, ob „fremde“ ökonomische Evidenz auf das eigene Land übertragen werden kann [14]. Auch aus anderen Gründen können Studienresultate für den Entscheidungskontext, in dem sie verwendet werden sollen, unterschiedlich brauchbar sein [15]. So kann eine Versichertengemeinschaft eine andere Altersstruktur – und damit andere Kosten und gesundheitliche Ergebnisse – aufweisen als die zugrunde gelegte Studie, oder die Erfahrung der tatsächlich tätigen Ärzte in der Durchführung neuer Eingriffe kann von derjenigen der Studienärzte abweichen und dann zu vermehrten Komplikationen führen. Auch die Mitwirkung von Patienten und Ärz-

ten bei der Versorgung kann variieren [16]. Manche dieser Unterschiede lassen sich herausrechnen, bei anderen muss eine neue Studie durchgeführt werden. In jedem Fall ist die Brauchbarkeit einer Studie für einen Entscheidungskontext zu prüfen.

Wegen der besonderen Rolle des Entscheidungskontexts können bei der ökonomischen Evidenz auch nicht die gleichen Evidenzmaßstäbe angelegt werden wie bei der medizinischen Effektivität. So wird bei ökonomischen Studien eine starke Evidenz nicht dann erreicht, wenn mehrere qualitativ hochwertige Studien unterschiedlicher nationaler oder internationaler Forschergruppen zum gleichen Ergebnis führen, sondern dann, wenn eine oder mehrere hochwertige Studien möglichst präzise auf den zu unterstützenden Entscheidungszusammenhang ausgerichtet sind. Auch können geringere Evidenzgrade in der Ökonomie nicht so leicht voneinander abgegrenzt werden. So lässt sich die Evidenz auf der Basis einer gut auf den Entscheidungskontext passenden Beobachtungsstudie schwer vergleichen mit einer weniger gut passenden, aber methodisch hochwertigen prospektiven klinischen Versuchsstudie. Da der Entscheidungskontext, in dem die Information verwendet werden soll, wesentlich ist, muss die ökonomische Evidenz unter mehr Aspekten geprüft werden als die medizinische Evidenz, und internationale Standards sind nur eingeschränkt möglich.

17.11 Kann ökonomische Evidenz helfen, den Leistungskatalog zu gestalten?

Am Ende des Ausgangszenarios wird die Nutzung der ökonomischen Evidenz bei der Bestimmung des Leistungskatalogs einer Krankenversicherung angesprochen. Eine im angelsächsischen Raum viel diskutierte Idee geht von einem Gesundheitssystem aus, in dem ein festes Budget für Gesundheitsleis-

tungen zur Verfügung steht und in dem die Entscheidungsträger das Ziel haben, die Gesundheit der Bevölkerung – messbar an den QALYs – zu maximieren. Dazu müssten alle möglichen medizinischen Maßnahmen ökonomisch evaluiert und nach ihrer Wirtschaftlichkeit, gemessen an den Kosten pro gewonnenem QALY, in eine Rangfolge gebracht werden. Zunächst wäre die kostengünstigste Intervention zu finanzieren, dann die nächste und so fort, bis das Budget aufgebraucht ist. Normativ geht diese Idee auf den Utilitarismus zurück. Eine genauere Betrachtung zeigt grundlegende Probleme: Es erscheint kaum erreichbar, für alle Interventionen qualitativ hochwertige, methodisch vergleichbare und aktuelle Wirtschaftlichkeitsdaten zu erlangen. Ferner wird in unseren Gesellschaften nicht nur nach Wirtschaftlichkeit, sondern auch nach ethischen, juristischen und anderen Gesichtspunkten entschieden. Das theoretische Referenzkonzept ist somit weder realistisch noch wertkonform. Eine Verbesserung der Wirtschaftlichkeit medizinischer Versorgung erfordert aber keineswegs, dass eine Entscheidungsinstanz zugleich über alle möglichen Interventionen befindet. Dies hat im britischen Nationalen Gesundheitsdienst 1999 zum Aufbau des bereits erwähnten, für seine evidenzbasierte Arbeit sehr anerkannten National Institute for Health and Clinical Excellence [NICE] (www.nice.org.uk) geführt, das seine sukzessiven Empfehlungen zum Leistungskatalog routinemäßig auch auf die Beurteilung der Wirtschaftlichkeit stützt. Ohne den englischen Entscheidungsansatz zu werten, lässt sich feststellen, dass NICE mit der Transparenz seines Verfahrens und mit der wissenschaftlichen Fundierung seiner Diskussionen weltweit einen Standard gesetzt hat.

Auch im deutschen Gesundheitssystem bieten die Aufgaben des Gemeinsamen Bundesausschusses (www.g-ba.de) einen Ansatzpunkt zur Verwendung der ökonomischen Evidenz. Dieser Ausschuss hat bei seinen

Entscheidungen über die Aufnahme von medizinischen Interventionen in den Leistungskatalog der gesetzlichen Krankenversicherung gemäß § 135 des Sozialgesetzbuchs V die medizinische Effektivität und Wirtschaftlichkeit zu berücksichtigen. Dies kann auf der Basis ökonomischer Evidenz erfolgen. Allerdings kann sektorales Entscheiden das Ausnutzen von Wirtschaftlichkeitsspielräumen erschweren, wenn etwa teure und effektive Medikamente Arzneimittelbudgets belasten, Krankenhausbudgets aber entlasten und der Kostenvorteil nicht ausgeglichen werden kann. Das 2004 gegründete Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen [IQWiG] (www.iqwig.de) unterstützt den Gemeinsamen Bundesausschuss durch Berichte zur jeweils geforderten Evidenz, wobei die ökonomische Evidenz bei der Arzneimittelbewertung mit der Gesundheitsreform 2007 eine substantielle Rolle erhielt.

17.12 Welche anderen Entscheidungen kann ökonomische Evidenz unterstützen?

Die Nutzung einer Technologie hängt nicht allein von der Aufnahme in den Leistungskatalog der Krankenversicherung ab. So müssen etwa ein Krankenhausarzt, der unter einer Vergütung nach Fallpauschalen arbeitet, oder ein niedergelassener Arzt, der Wirtschaftlichkeitsprüfungen ausgesetzt ist, entscheiden, ob sie die im Anfangsszenario genannte Kombination aus DNA-Chip und teurer Chemotherapie bei ihren Patienten einsetzen können. Bereits vor einigen Jahren stellten O'Brien et al. [17] erstmals die Entwicklung einer Behandlungspolitik in einer Klinik unter Einschluss der ökonomischen Evidenz dar, und zwar am Beispiel der Wahl einer thrombolytischen Therapie nach akutem Herzinfarkt. Auch bei anderen Entscheidungen der Betriebsführung im Krankenhaus

oder in der Arztpraxis könnte die Wirtschaftlichkeit eine Rolle spielen, etwa bei der Bereitstellung von diagnostischen Kapazitäten. Wenn für neue Technologien die Wirtschaftlichkeit nachgewiesen wird, könnten umgekehrt Fallpauschalen und Gebührenordnungen angepasst oder Arzneimitteln von der Einstufung in Festbetragsgruppen ausgenommen werden. Anbieter von Gesundheitsleistungen können somit die ökonomische Evidenz in verschiedener Weise nutzen; Abbildung 17.3 zeigt für Ärzte, Krankenhäuser und die pharmazeutische Industrie vier entsprechende Entscheidungsbereiche.

Auch Behandlungsleitlinien lassen sich durch Informationen zur Wirtschaftlichkeit der empfohlenen Vorgehensweisen ergänzen. Die Daten zur Wirtschaftlichkeit geben noch nicht an, ob die für einen Gesundheitsgewinn erforderlichen zusätzlichen Kosten auch aufgewendet werden sollen; dies muss der verantwortliche Therapeut selbst abwägen. Daher kann eine Leitlinie nicht generell festlegen, was als wirtschaftlich zu gelten hat und durchzuführen wäre. Die Funktion der ökonomischen Evidenz in Leitlinien liegt darin, den Nutzer zu informieren und seine Entscheidungen zu fundieren. Als Instrument rationaler Entscheidungsunterstützung sollten Leitlinien die ökonomische Dimension medizinischen Handelns nicht ignorieren, sondern evidenzbasiert berichten.

Ökonomische Evidenz kann ferner Forschung orientieren, sei es bei der Definition neuer Vorhaben, mit denen auch die Wirtschaftlichkeit verbessert werden soll, oder bei Entscheidungen über die Fortführung laufender Vorhaben. Da viele Länder heute schon die Wirtschaftlichkeit bei der Vergütung neuer Arzneimittel berücksichtigen, spielt die ökonomische Evidenz in der Entwicklung und Vermarktung neuer Präparate eine wichtige Rolle. Sie wird ergänzend zur medizinischen Wirksamkeit eingesetzt, um eine adäquate Vergütung zu erzielen, im Produktmarketing verwendet und in frühen

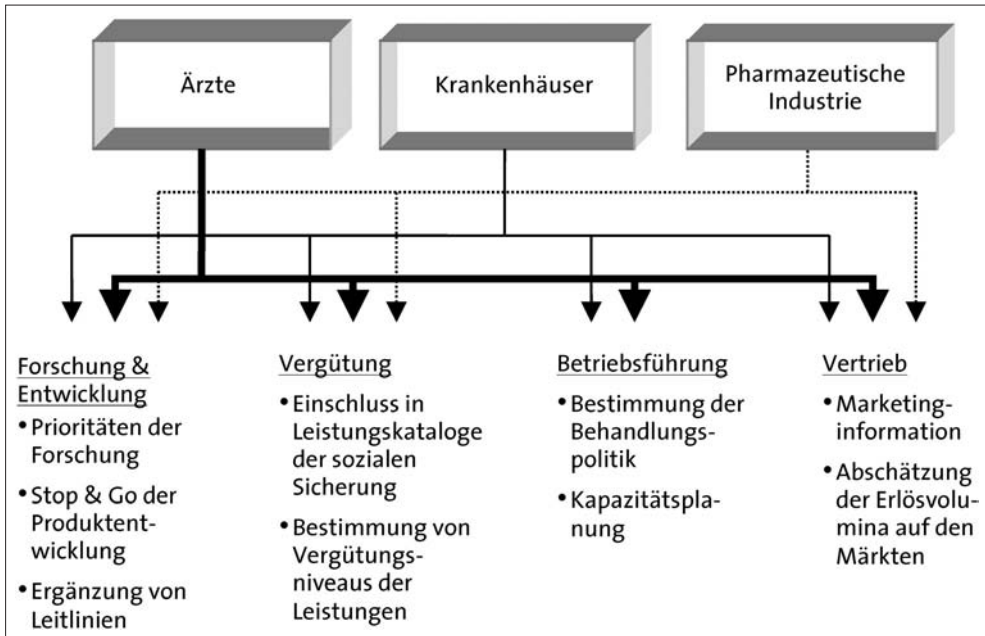


Abb. 17.3: Nutzung ökonomischer Evidenz durch Leistungsanbieter

Forschungsphasen zur Abschätzung der Marktfähigkeit der Produkte gebraucht.

17.13 Ausblick und Auflösung des Szenarios

In einem rational gestalteten Wettbewerb um knappe Mittel im Gesundheitswesen können sich diejenigen Akteure besser behaupten, die zeigen können, dass ihre Leistungen und Maßnahmen effektiv und kosteneffektiv sind. Zur evidenzbasierten Verbesserung der Effizienz muss hochwertige Information zielkonform genutzt werden: Die Evidenz sollte möglichst aktuell, umfassend und qualitativ akzeptabel sein, der Entscheidungskontext sollte klar definiert sein und Anreize zur Wirtschaftlichkeit setzen. Die Nutzung ökonomischer Evidenz sollte innerhalb eines klaren normativen Rahmens erfolgen, der eine Abstimmung mit anderen Entscheidungskriterien wie ethischen oder juristischen Gesichtspunkten vorsieht; für die Beteiligten und Betroffenen sollten die

Kriterien und Prozesse transparent sein. In der Realität wird es Abstriche gegenüber diesem Idealbild geben. Eine Entscheidung ohne Berücksichtigung der ökonomischen Evidenz bietet aber rational keine bessere Alternative.

Auflösung des Szenarios

Im anfangs beschriebenen Szenario sehen schließlich die Klinikchefin und der Krankenkassenvorstand, dass selbst eine hochwertige amerikanische Studie über den Therapieeinsatz ohne Chip die ökonomischen Auswirkungen der neuen Vorgehensweise für ein europäisches Land nicht adäquat erfassen kann. Verglichen werden müssen die zusätzlichen Kosten und Gesundheitseffekte der Behandlung der Patientengruppe nach DNA-Screening, und zwar unter den Bedingungen des eigenen Gesundheitssystems und nach den dort gültigen Richtlinien für die ökonomische Evaluation. Wenn geeignete Daten vorliegen, könnte dies auch durch eine Modellierung abgeschätzt werden. Die Klinikchefin will ihrer Fachgesellschaft eine ökonomisch begleitete klinische Studie vorschlagen und die Ergebnisse in

die Leitliniendiskussion einbeziehen. Der Kassenvorstand beschließt, die Kosten pro QALY in einem Modell intern ermitteln zu lassen und die Resultate in die nächste Diskussion über Arzneimittelpreise einzubringen.

Literatur

- [1] Leidl R, von der Schulenburg JM, Wasem J (1999) Ansätze und Methoden der ökonomischen Evaluation – eine internationale Perspektive. Nomos, Baden-Baden
- [2] Drummond MF et al. (2005) Methods for the economic evaluation of health care programmes. Oxford Univ. Pr., Oxford
- [3] Leidl R (2003) Der Effizienz auf der Spur: eine Einführung in die ökonomische Evaluation. In: Schwartz FW et al., Das Public Health Buch. Gesundheit und Gesundheitswesen, 461–484. Urban und Fischer, München
- [4] Schöffski O, von der Schulenburg JM, (Eds) (2002) Gesundheitökonomische Evaluationen. Grundlagen und Standortbestimmung. Studienausgabe. Springer, Berlin
- [5] Gold MR et al. (1996) Cost-effectiveness in health and medicine. Oxford Univ. Pr., Oxford
- [6] Drummond MF, Jefferson TO, Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. (1996), 313, 275–283
- [7] Philips Z et al., Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. (2006), 24, 355–371
- [8] Evers S et al., Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. Int J Technol Assess Health Care. (2005), 21, 240–245
- [9] Krauth C et al., Empirische Bewertungssätze in der gesundheitsökonomischen Evaluation – ein Vorschlag der AG Methoden der gesundheitsökonomischen Evaluation (AG MEG). Gesundheitswesen. (2005), 67, 736–746
- [10] Hill SR, Mitchell AS, Henry DA, Problems with the interpretation of pharmacoeconomic analyses: a review of submissions to the Australian Pharmaceutical Benefits Scheme. JAMA. (2000), 283, 2116–2121
- [11] Goossens ME et al., The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. J Clin Epidemiol. (2000), 53, 688–695
- [12] Leidl R et al., Do physicians accept quality of life and utility measurement? An Austrian, Finnish, and German survey. Health Econ Prev Care. (2001), 2, 170–175
- [13] Briggs AH, O'Brien BJ, Blackhouse G, Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. Annu Rev Public Health. (2002), 23, 377–401
- [14] Welte R et al., A decision chart for assessing and improving the transferability of economic evaluation results between countries. Pharmacoeconomics. (2004), 22, 857–876
- [15] Baltussen R, Leidl R, Ament A, Real world designs in economic evaluation. Bridging the gap between clinical research and policy-making. Pharmacoeconomics. (1999), 16, 449–458
- [16] Tugwell P et al., The measurement iterative loop: a framework for the critical appraisal of need, benefits and costs of health interventions. J Chronic Dis. (1985), 38, 339–351
- [17] O'Brien BJ et al., Users' guides to the medical literature. XIII. How to use an article on economic analysis of clinical practice. B. What are the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. JAMA. (1997), 277, 1802–1806

18 Kritische Bewertung von Gesundheitsinformationen für medizinische Laien

Sylvia Sanger, Marie-Luise Dierks

18.1 Hintergrund

Das Ziel der Arzt-Patienten-Kommunikation ist es unter anderem, auf der Basis geteilter Informationen zu einer gemeinsam verantworteten Entscheidung in Bezug auf Diagnostik und Therapie zu gelangen [1]. Evidenzbasierte Informationen fur Arzte und Patienten nehmen somit im Rahmen der gemeinsamen Entscheidungsfindung (Shared decision making) eine Schlusselrolle ein. Sie dienen neben der Entscheidungsfindung uber eine angemessene Vorgehensweise im individuellen Fall vor allem einer Qualitatsverbesserung der medizinischen Versorgung [2].

Das Informationsverhalten der Patienten hat in den vergangenen Jahren eine Trendwende erfahren, die sich zukunftig noch starker auswirken wird. Bisher wurde der Arzt als wichtigste Instanz der Informationsvermittlung im Bereich von Gesundheit und Krankheit [3] angesehen. Moderne Medien – hier ist insbesondere das Internet zu nennen – versorgen nun Patienten mit einer Fulle an Informationen. Zudem haben sich Gesundheitssendungen in Rundfunk und Fernsehen einen Markt erobert [4].

Immer haufiger kommen Patienten deshalb mit einer gewissen Vorerfahrung und vor allem Vorinformation in die Arztpraxis, die sie aus den genannten Medien bezogen haben, und sie informieren sich parallel und erganzend zum Therapieverlauf uber ganz unterschiedliche Quellen, vom Internet angefangen uber diverse Printmedien bis hin zu Beratungs-Hotlines der Krankenkassen oder unabhangiger Patientenberatungsstel-

len [5–7]. Dabei stellt sich grundsatzlich die Frage, welche Qualitat die vermittelten Informationen haben und welche Moglichkeiten fur Professionelle und Laien existieren, die wissenschaftliche Evidenz und Glaubwurdigkeit von Informationen zu uberprufen. Dies gilt fur die Inhalte von Beratungsgesprachen ebenso wie fur schriftliche Informationen und Internetseiten [8].

Arzte und Patienten brauchen eine einheitliche und insbesondere fur Patienten verstandliche wissenschaftliche Informationsbasis. Wenn Informationen, die Arzte und Patienten ihren Entscheidungen zugrunde legen, inhaltlich voneinander abweichen, kann das zu Kommunikationsstorungen und letztlich einem Vertrauensverlust in der Arzt-Patienten-Beziehung fuhren. Der Prozess der Informationsbeschaffung stellt hohe Anforderungen an Patienten. Sie mussen diese in den entsprechenden Medien mittels geeigneter Suchstrategien erst einmal finden und dann auch noch auf ihre Glaubwurdigkeit und Relevanz hin beurteilen. Was schon fur Professionelle, die in der Regel uber Kenntnisse in medizinischer Literaturrecherche verfugen sollten, zum Problem werden kann (nur 3% von 2543 befragten Arzten zeigten sich zufrieden mit ihrer Situation bei der Informationsbeschaffung [9]), scheint fur Laien kaum zu bewaltigen. Trotz des nahezu unuberschaubaren Angebotes an medizinischen Informationen sowohl im Printbereich als auch im Internet gibt es im deutschsprachigen Raum bisher wenige Angebote, die wissenschaftlich begrundete (evidenzbasierte) Ratschlage fur medizinische Laien aufbereitet enthalten. Patienten und

Verbraucher können in der Regel nicht einschätzen, ob sich eine Information tatsächlich auf die derzeit best verfügbare Evidenz stützt. Hinzu kommt, dass sie entsprechende Konzepte weder kennen noch tatsächlich anwenden können.

Das vorliegende Kapitel beschreibt, was Sie als Ärztin/Arzt unter Berücksichtigung Ihres Zeitbudgets tun können, um in Ergänzung zum Arzt-Patienten-Gespräch verlässliche Informationen aus und für Patientensicht zu identifizieren und Ihre Patienten mit wenig Zeitaufwand effektiv bei der Suche nach evidenzbasierten Patientinformationen zu unterstützen. Der Beitrag enthält auch Hinweise darauf, wo Sie Fachinformationen für Professionelle finden können. Im Kapitel 18.4 „Werkzeugkasten für Arzt und Patient“ haben wir eine Reihe von praktischen Hilfen, Informations- und Kopiervorlagen zusammengestellt, die Sie im Anhang des Buches finden. Diese Materialien stehen auch auf den Internetseiten des Deutschen Netzwerk Evidenzbasierte Medizin e.V. (www.ebm-netzwerk.de/grundlagen/arbeitsmaterialien) zum Download zur Verfügung.

18.2 Szenarien

Nachfolgend sollen drei Beispiele aus dem klinischen Alltag besprochen und der Weg der Informationsgewinnung und kritischen Bewertung der Informationen aus und für Patientensicht sowie Aspekte der Kommunikation aufgezeigt werden. Es handelt sich dabei um:

▲ **Szenario 1:** Ein häufig auftretendes gesundheitliches Problem ohne akuten Handlungsbedarf für den Arzt mit guter Evidenzlage und ausreichend Patientinformationen. Hier wurde die gut bekannte Diskussion um die Hormonersatz-Therapie zur Behandlung der Wechseljahresbeschwerden aufgegriffen, um exemplarisch darzustellen, wie Patienten bei einer Fra-

ge, zu der umfassende Informationen vorliegen, für sie relevante Informationen finden und bewerten können.

- ▲ **Szenario 2:** Ein häufig auftretendes gesundheitliches Problem mit akutem Handlungsbedarf für den Arzt bei guter Datenlage und wenig Patientinformationen. Hier wurde die Gabe von nicht-steroidalen Antiphlogistika intramuskulär bei Rückenschmerz gewählt.
- ▲ **Szenario 3:** Ein selten auftretendes gesundheitliches Problem mit akutem Handlungsbedarf für den Arzt bei schlechter Evidenzlage und kaum verfügbaren Patientinformationen. Das dritte Beispiel stellt die Situation bei einer seltenen Erkrankung am Beispiel Chorea Huntington dar.

18.2.1 Szenario 1

*Stellen Sie sich vor, Sie sind Gynäkologe/in und beginnen Ihren Arbeitstag mit einem vollen Wartezimmer. Eine der Patientinnen, eine resolute Endvierzigerin, klagt über Beschwerden, wie sie typischerweise beim Einsetzen der Wechseljahre auftreten. Die Patientin hat sich bereits ausführlich im Internet über ihre Beschwerden informiert. Sie hat bei der Suche nach den Stichwörtern „**beschwerden in den wechseljahren**“ unter anderem die Seiten eines kommerziellen Anbieters über die Hormontherapie (z.B. <http://www.solvay-arzneimittel.de/Patient/Gynaekologie/Beschwerden/>) gefunden. Den Internetausdruck hat sie dabei. Die Patientin ist nach der Lektüre vom Nutzen dieser Therapie überzeugt (von möglichen Gefahren oder Schäden ist in der entsprechenden Information keine Rede) und möchte nun eine entsprechende Medikation erhalten. Sie selbst kennen auch die kritischen Studienergebnisse und weisen die Patientin darauf hin, dass eine Hormontherapie potenziellen Schaden verursachen kann. Die Patientin beharrt jedoch auf dieser Behandlungsoption, zumal ihr eine Bekannte, der sie sehr vertraut, diese Empfehlung*

ebenfalls gegeben hat. Sie weist darauf hin, dass in dem Internetausdruck, der ja von einer seriösen Pharmafirma stamme, von möglichen Schäden der Hormontherapie keine Rede sei. Außerdem habe sie in einem Frauenmagazin gelesen, dass viele Schauspielerinnen auch Hormone nahmen und sich gut damit fühlten.

Sie stehen nun vor den Aufgaben:

- ▲ Der Patientin zu erläutern, welche Möglichkeiten sie hat, ihre Beschwerden zu lindern und worin der Nutzen und die Nebenwirkungen der einzelnen Möglichkeiten bestehen.
- ▲ Gemeinsam mit der Patientin zu entscheiden, welcher Behandlungsweg infrage kommt.

Wo liegt das Problem?

Durch die Internetrecherche schon vor dem Arztbesuch und durch die Gespräche mit ihrer Freundin hat sich die Patientin bereits eine klare Meinung gebildet. Einiges von dem, was Sie sagen oder empfehlen, könnte dieser Meinung widersprechen und sich vielleicht sogar negativ auf Ihr gemeinsames Vertrauensverhältnis auswirken. Alles sieht nach einer langen und vielleicht wenig erbaulichen Diskussion aus, und Sie wissen, Ihr Wartezimmer ist voll! Sie gehen nach wie vor davon aus, dass sich die Patientin nur dann in einer guten Weise für eine Behandlung (oder Nichtbehandlung) entscheiden kann, wenn sie alle Optionen kennt und deren Nutzen und mögliche Nebenwirkungen abwägen kann.

Möglicher Lösungsweg

Ihr Ziel ist es:

1. sich selbst aktuelle ärztliche Fachinformationen über dieses Thema zu beschaffen,
2. die Patientin auf verlässliche medizinische Laieninformationen zu diesem Thema hinzuweisen (dazu gehört auch, ihr zu erklären, woran man verlässliche Informationen erkennen kann),

3. mit der Patientin die eingangs mitgebrachte Patienteninformation zu besprechen,
4. und auf dieser Basis eine gemeinsam verantwortete Entscheidung zu treffen.

Dies kann bedeuten, dass Sie mit der Patientin mindestens einen weiteren Termin vereinbaren müssen. Das mag auf den ersten Blick überzogen erscheinen. Jedoch schafft eine solche Vorgehensweise eine gute Vertrauensbasis zwischen Arzt und Patient, auf deren Grundlage dann letztlich eine gemeinsam verantwortete Entscheidung getroffen werden kann.

18.2.1.1 Beschaffung aktueller ärztlicher Fachinformationen über das Thema Hormonersatz-Therapie in den Wechseljahren

Zunächst eine kurze Definition: Eine Information ist dann evidenzbasiert, wenn Aussagen zu Untersuchungen und Behandlungsmöglichkeiten mit den wissenschaftlichen Quellen belegt sind, die die zum Zeitpunkt der Erstellung besten und aussagekräftigsten Daten zum betreffenden Problem beinhalten. Die Datenlage ist im besten Fall eine systematische Übersichtsarbeit mit Meta-Analyse oder eine randomisierte kontrollierte klinische Studie. Die am wenigsten gesicherte Datenlage ist die konsentiertere Meinung von in dem Gebiet spezialisierten Fachleuten (Experten), da sie auf subjektiven Einzelerfahrungen beruht [10]. Es gilt also, zur interessierenden Frage die bestmögliche Evidenz zu finden, die in manchen Fällen – insbesondere bei sehr seltenen Erkrankungen – durchaus die konsentiertere Meinung von Experten sein kann.

Um entsprechende Informationen zu finden, können Sie zunächst in einer großen Datenbank mit zahlreichen systematischen Übersichtsarbeiten suchen. In der Cochrane Library finden Sie, bezogen auf das genannte Beispiel, eine systematische Übersichtsarbeit

zum Titel „Long term hormone therapy for perimenopausal and postmenopausal women“. Hier werden bereits detailliert Chancen und Risiken aufgelistet.

In einem weiteren Schritt können Sie in eine Suchmaschine (z.B. www.google.de) die Stichworte „**hormontherapie**“ und „**leitlinie**“ oder „**evidenzbasiert**“ eingeben. Dabei finden Sie unter anderem die Therapieempfehlung „Hormontherapie im Klimakterium“ der Arzneimittelkommission der Deutschen Ärzteschaft (AkdÄ). Sie suchen in der Leitliniendatenbank der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF; <http://leitlinien.net/>) unter der Rubrik „Gynäkologie und Geburtshilfe“ und finden die „Konsensusempfehlungen zur Hormontherapie (HT) im Klimakterium und in der Postmenopause“, die von der Deutschen Gesellschaft für Gynäkologie und Geburtshilfe (DGGG) im Konsens mit dem

Berufsverband der Frauenärzte erstellt wurde (<http://www.uni-duesseldorf.de/AWMF/II/015-038.htm>). Diese Leitlinie ist gekennzeichnet als so genannte S1-Leitlinie, d.h. dass ihr keine formale Evidenzrecherche zugrunde liegt und sie nicht systematisch entwickelt wurde. In der Quellenangabe finden Sie jedoch randomisierte kontrollierte Studien (RCTs) und Meta-Analysen zum Thema. In beiden ärztlichen Informationen sind die Vor- und Nachteile einer Hormonersatz-Therapie dargestellt.

18.2.1.2 Information der Patientin über verlässliche medizinische Laieninformationen zu diesem Thema

Da es ein zentrales, unabhängiges Patientenportal mit verständlichen Informationen zu allen Erkrankungsgebieten in Deutschland noch nicht gibt – derzeit baut das Institut für Qualität und Wirtschaftlichkeit im Gesund-

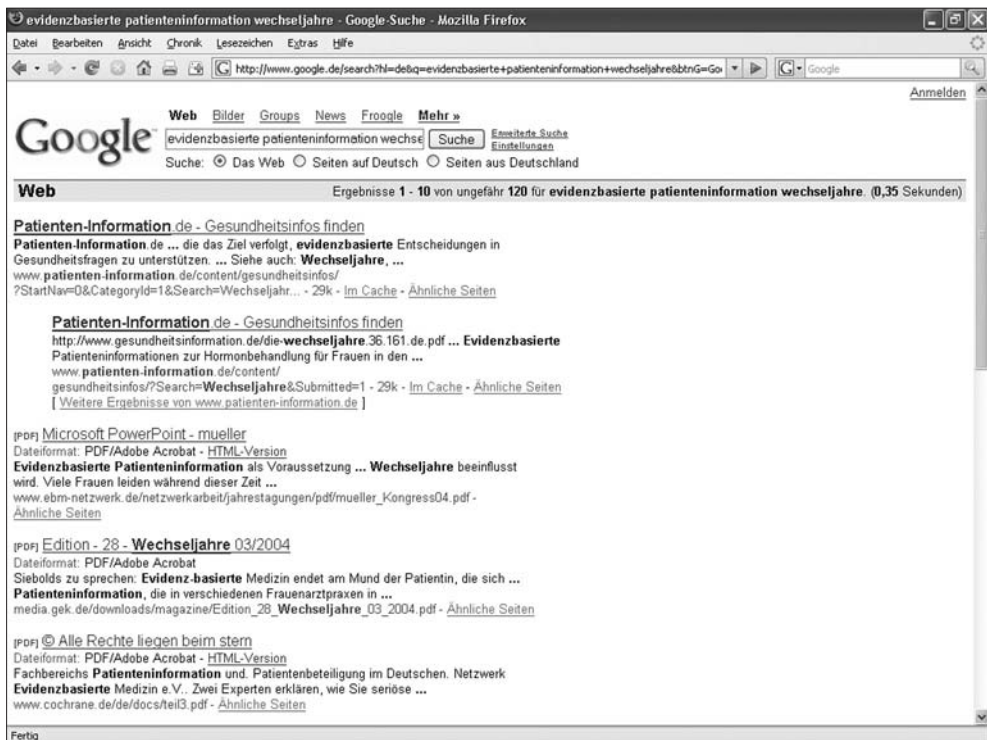


Abb. 18.1: Google-Suche nach „evidenzbasierte patienteninformation wechseljahre“

heitswesen (IQWiG) unter der Adresse www.gesundheitsinformation.de ein Patientenportal mit evidenzbasierten Informationen auf, das Angebot wird ständig erweitert –, benutzen Sie eine Suchmaschine wie z.B. Google. Hier geben Sie die Begriffe „**evidenzbasierte patienteninformation wechseljahre**“ ein (s. Abb. 18.1). Bei der Suche kann es übrigens hilfreich sein, unterschiedliche Schreibweisen einzugeben und die erweiterte Suchfunktion zu nutzen. Ein Blick in die Online-Hilfe von Suchmaschinen hilft oft, die richtigen Schreibweisen bei der Verknüpfung der Suchbegriffe anzuwenden. Im Suchergebnis zum verwendeten Begriff stoßen Sie unter anderem auf den Patienten-Informationssdienst der Deutschen Ärzteschaft (www.patienten-information.de).

Der Patienten-Informationssdienst der Deutschen Ärzteschaft bietet Ihnen Links zu drei Patienteninformationen zu Wechseljah-

ren an. Die Anbieter dieser Informationen sind das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, die Arzneimittelkommission der deutschen Ärzteschaft, auf deren ärztliche Leitlinie Sie bereits gestoßen sind, und die Fachrichtung Gesundheit des Instituts für Gewerblich-Technische Wissenschaften der Universität Hamburg (s. Abb. 18.2).

Für diese Patienteninformationen (s. Abb. 18.2), wurde bereits eine Bewertung der methodischen Qualität mit dem DISCERN-Instrument [11] vorgenommen (s. Tab. 18.1). Die Emoticons („Smileys“) am rechten Seitenrand zeigen Ihnen auf einen Blick, dass alle diese Informationen von den Bewertungsexperten als gut bis sehr gut eingeschätzt wurden. Da die Patientin eine eifrige Internetnutzerin ist, geben Sie ihr die entsprechenden Links an die Hand, damit sie zu Hause in Ruhe die Informationen nachlesen kann.

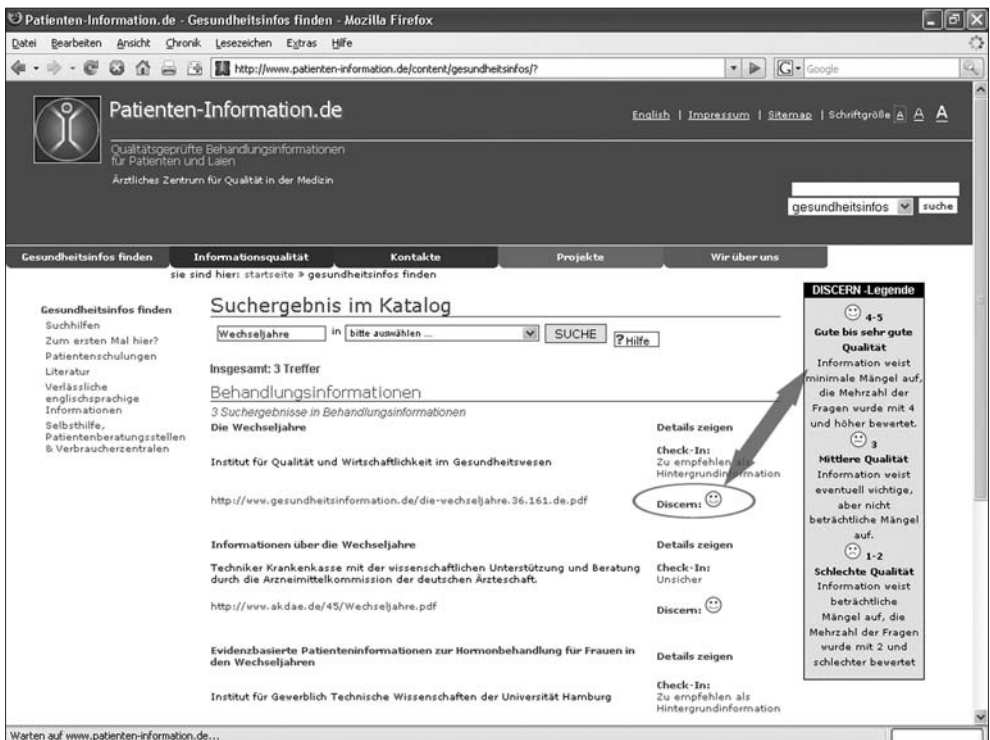


Abb. 18.2: Evidenzbasierte Patienteninformationen zur Hormontherapie in den Wechseljahren, gelistet in www.patienten-information.de

Falls Sie die technischen Möglichkeiten haben und in Ihrer Praxis offensiv nutzen wollen, wäre es auch möglich, den Internetzugang als Informationstool – gegebenenfalls zusammen mit einschlägigen Checklisten und Hinweisen versehen (s. auch Vorlagen unter der Rubrik „Werkzeugkasten für Arzt und Patient“) – direkt in den Praxisräumen zu ermöglichen und von hier aus mit den Patienten das Gespräch zu suchen.

18.2.1.3 Besprechung der von der Patientin eingangs mitgebrachten Patienteninformation

Sie haben jetzt die nötigen Informationen, um Ihre Patientin zu beraten, und auch die Patientin hat einschlägige Quellen gefunden. Sie müssen nun mit ihr die eingangs mitgebrachte Information besprechen und ihr die Möglichkeit geben, diese mit den von Ihnen recherchierten Informationen zu vergleichen. Um dies sachlich untermauert tun zu können, sollten Sie Kriterien zur formalen Überprüfung der Qualität von Patienteninformationen verwenden. Bei der Beurteilung der methodischen Qualität von Patientenin-

formationen hat sich unter zahlreichen Instrumenten [12] das validierte DISCERN-Instrument [10] (s. Tab. 18.1) bewährt, das auch im internationalen Raum in den vergangenen Jahren zunehmend Verbreitung gefunden hat. Wie in Abbildung 18.2 dargestellt, sind beispielsweise die Behandlungsinformationen, die Sie in der kommentierten Linksammlung des Patienten-Informationsdienstes der Ärzteschaft gefunden haben, nach diesen Kriterien bewertet worden.

Die Kriterien des DISCERN-Instruments (s. Tab. 18.1) bieten Ihnen einen Anhaltspunkt für die Diskussion der von der Patientin mitgebrachten Materialien, aber auch für die in der neuen Recherche gefundenen Patienteninformationen.

Sie können zwar mit den 15 Fragen des Instrumentes DISCERN nicht abschließend bewerten, inwieweit tatsächlich die best verfügbare Evidenz dargestellt wird (dies ist mit keinem Fragespiegel möglich, sondern kann letztlich nur von auf dem jeweiligen Fachgebiet besonders qualifizierten Experten beurteilt werden), jedoch können Sie auf dieser Basis wichtige Qualitätsaspekte diskutieren. Wenn

Tab. 18.1: Die Kriterien des DISCERN-Instruments

Zuverlässigkeit der Informationen:
1. Sind die Ziele der Publikation klar?
2. Erreicht die Publikation die selbst gesteckten Ziele?
3. Ist die Publikation für Sie bedeutsam?
4. Existieren klare Angaben zu den Informationsquellen, die zur Erstellung herangezogen wurden?
5. Ist klar angegeben, wann die Informationen der Publikation erstellt wurden?
6. Ist die Publikation ausgewogen und unbeeinflusst?
7. Enthält die Publikation detaillierte Angaben über ergänzende Hilfen und Informationen?
8. Äußert sich die Publikation zu Bereichen, für die keine sicheren Informationen vorliegen?
Qualität der Info zu Behandlungsalternativen:
9. Beschreibt die Publikation die Wirkungsweise jedes Behandlungsverfahrens?
10. Beschreibt die Publikation den Nutzen jedes Behandlungsverfahrens?
11. Beschreibt die Publikation die Risiken jedes Behandlungsverfahrens?
12. Beschreibt die Publikation mögliche Folgen einer Nichtbehandlung?
13. Beschreibt die Publikation, wie die Behandlungsverfahren die Lebensqualität beeinflussen?
14. Ist klar dargestellt, dass mehr als ein mögliches Behandlungsverfahren existieren kann?
15. Ist die Publikation eine Hilfe für eine „partnerschaftliche Entscheidungsfindung“?

bereits die genannten, eher formalen Anforderungen nicht erfüllt sind, handelt es sich ohnehin nicht um eine evidenzbasierte Information, und der Leser sollte in jedem Fall weitere Informationsquellen heranziehen – das gilt im Übrigen auch für persönliche Beratungsgespräche, die auf möglichst evidenzbasierten Hintergrundinformationen basieren sollten.

Am Beispiel ausgewählter Indikatorfragen und der Einschätzung der Qualität der von der Patientin mitgebrachten Information zur Hormontherapie (Informationsseiten der Pharmaindustrie) wird nachfolgend die Relevanz des Instrumentes deutlich. Diese Indikatorfragen werden im übrigen in den verschiedensten Instrumenten zur Qualitätsbewertung von Informationen ebenfalls genutzt (so z.B. bei www.hon.ch, www.afgis.de, www.quick.org.uk).

Ist klar angegeben, wann die Information erstellt wurde?

Anhand der von der Patientin vorgelegten Internetinformation lässt sich keine Aussage zur Aktualität machen, es ist kein Erstellungsdatum angegeben. Entsprechend ist hier erforderlich, noch genauer nach dem Erstellungsdatum der Information zu suchen oder gegebenenfalls andere (und in diesem Falle aktuellere) Daten heranzuziehen.

Existieren klare Angaben zu den Informationsquellen, die zur Erstellung herangezogen wurden?

Sie sehen, dass weder ein Autor noch etwaige verwendete Quellen angegeben sind. Das heißt, es werden keine wissenschaftlichen Belege über die Richtigkeit der in der Information gemachten Aussagen vorgelegt, eine entsprechende Überprüfung der Aussagen wird so unmöglich.

Ist die Publikation ausgewogen und uneinflusst?

Es handelt sich um eine Information eines Arzneimittelherstellers. Deshalb muss sie

nicht per se schlecht sein. Doch sollte man bedenken, dass hier der Verkauf von Produkten im Vordergrund der Patienteninformation stehen könnte, entsprechend wäre die Nutzung zusätzlicher Informationsmaterialien angezeigt.

Sind Wirkungsweise, Nutzen, Risiken der Therapie(n) und mögliche Folgen einer Nichtbehandlung dargelegt?

Wird beschrieben, wie sich die Behandlung auf die Lebensqualität auswirkt?

Ist dargestellt, dass (ob) mehr als ein mögliches Behandlungsverfahren infrage kommt?

Es wird zwar erwähnt, dass die Beschwerden auch ohne eine Behandlung mit den Jahren wieder nachlassen und sogar vollständig verschwinden können, jedoch finden Sie keine Aussagen zu möglichen Risiken der Behandlung oder wie sich diese auf die Lebensqualität auswirkt – außer dem Versprechen, dass eine Besserung der Befindlichkeit erreicht werde. Alternativen einer Behandlung werden nur insofern dargestellt, als auf unterschiedliche Hormonkombinationen hingewiesen wird.

Zusammenfassend ergibt schon diese Kontrastierung der Patienteninformation mit den zentralen Qualitätsfragen, dass die Patientin nicht auf alle wichtigen Fragen eine Antwort bekommen hat und zusätzliche Informationsquellen braucht, um eine informierte Entscheidung treffen zu können.

18.2.1.4 Treffen einer gemeinsam verantworteten Entscheidung

Sie haben sich über die aktuelle Datenlage zum Thema Hormontherapie informiert. Sie haben Ihrer Patientin das Handwerkszeug vermittelt, geeignete Informationen zu identifizieren und ihr darüber hinaus verlässliche, evidenzbasierte Quellen genannt. Nun sind Sie beide auf dem gleichen Informationsstand – die Entscheidung steht an.

Möglicherweise müssen Sie in dieser Situation den Nutzen und die Risiken insbesondere der Hormonersatz-Therapie nochmals zusammenfassen. Dabei sollten Sie diese Informationen in einer Form angeben, die der Patientin eine realistische Einschätzung der Wirksamkeit der zur Diskussion stehenden Maßnahme ermöglicht. Am leichtesten verständlich ist die Darstellung mittels absoluter Zahlen [9; 13]. Achten Sie darauf, Fremdwörter zu erklären bzw. deutsche Wörter zu benutzen und erkundigen Sie sich immer wieder, ob die Patientin alles verstanden hat oder noch weitere Fragen hat. Formulierungen wie „Was müssen Sie noch wissen?“ oder „Welche Fragen bleiben noch?“ geben der Patientin die Gelegenheit, weiter nachzufragen und signalisieren auch, dass Sie selbst bereit sind, sich weiter auf das Gespräch einzulassen. Hilfreich für das Verständnis ist es auch, die Patientin mit eigenen Worten wiedergeben zu lassen, was sie wie verstanden hat.

18.2.2 Szenario 2

Ein Patient um die vierzig wird in Ihrer Allgemeinarztpraxis vorgestellt und klagt über starke Rückenschmerzen. Die, so sagt er, habe er schon des Öfteren gehabt. Eine Injektion mit Diclofenac habe ihm immer geholfen, das sei jedes Mal ganz schnell gegangen. Außerdem sei er beruflich stark gefordert und könne sich keine langen Ausfallzeiten erlauben, er möchte schnell wieder an die Arbeit zurück. Beiläufig erwähnt er, dass er eine überwiegend sitzende Bürotätigkeit ausübe.

Sie stehen nun vor den Aufgaben:

- ▲ Dem Patienten zu erläutern, dass es außer der „Spritze“ weitere Möglichkeiten gibt, seine Schmerzen zu lindern, und ihm den Nutzen sowie möglichen Schaden der Optionen darzulegen.
- ▲ Mit dem Patienten eine gemeinsam verantwortete Entscheidung zu treffen.

Wo liegt das Problem?

Es sitzt ein Patient vor Ihnen, der einen hohen Erwartungsdruck hat. Spritzen haben ihm immer geholfen, er verbindet damit eine schnelle Linderung seiner Beschwerden und möchte sich im Grunde auf keine andere Therapie einlassen. Sie würden eher zu einer oralen Medikamentengabe raten und den Patienten zudem ermutigen, sich sportlich zu betätigen. Aus Ihrer Sicht wäre eine orale Medikamentengabe besser geeignet als die Injektionen, die die Gefahren eines Abszesses, einer Gewebereizung oder der Verletzung eines Nervs bergen. Der Patient sagt Ihnen jedoch deutlich, dass ja bisher immer alles gut gegangen sei, wenn er eine Spritze bekommen hat. Ihre Meinungen bezüglich der Therapie sind also unterschiedlich.

Möglicher Lösungsweg

1. Beschaffung ärztlicher Informationen zur Therapie des Rückenschmerzes,
2. Abgleich der Präferenzen des Patienten mit den vorliegenden Evidenzen,
3. Treffen einer gemeinsam verantworteten Entscheidung,
4. Empfehlung von Informationsquellen für Laien an Ihren Patienten.

18.2.2.1 Beschaffung ärztlicher Informationen zum Thema Rückenschmerz

(S. auch Toolbox „Evidenzquellen für den Arzt“ im Anhang des Buches) Sie suchen nach ärztlichen Leitlinien zur Behandlung des Rückenschmerzes. Dazu geben Sie in die Suchmaschine Google (<http://www.google.de>) die Suchbegriffe „leitlinie rückenschmerz“ und „leitlinie kreuzschmerz“ ein. Sie stoßen auf die Clearing-Berichte des Ärztlichen Zentrums für Qualität in der Medizin (ÄZQ) zum akuten und chronischen Rückenschmerz (<http://www.leitlinien.de>), auf die Leitlinie Akuter Rückenschmerz der Deutschen Gesellschaft für Physikalische Medizin und Rehabilitation (DGPMR);

<http://www.uni-duesseldorf.de/AWMF/11/036-002.htm>), auf die Leitlinie Kreuzschmerzen der Deutschen Gesellschaft für Allgemein- und Familienmedizin (DEGAM; http://www.degam.de/leitlinien/3_kreuzschmerzen.html) und auf eine Fülle von weiteren Informationen. Um sich einen Überblick zu verschaffen, ob es systematische Übersichtsarbeiten zur Überlegenheit oraler Medikamente gegenüber intramuskulär verabreichter Medikamente gibt, schauen Sie zunächst in den kostenfrei zugänglichen Abstracts der Cochrane Library (<http://www.cochrane.de/de/browse.htm>) nach und finden 79 Zusammenfassungen zum Thema Rückenschmerz, von denen Sie drei für Ihre Frage („Welche Therapie ist wirkungsvoll?“) relevante Arbeiten finden. Eine weist darauf hin, dass es keine überzeugende Evidenz für die Überlegenheit der Injektionstherapie gibt. In der DEGAM-Leitlinie finden Sie sogar den Hinweis darauf, dass die intravenöse oder intramuskuläre Injektion von Schmerzmitteln, insbesondere von Diclofenac, zu den abzulehnenden Optionen gehört.

18.2.2.2 Abgleich der Präferenzen des Patienten mit den vorliegenden Evidenzen

Der Patient möchte die Inhalte der ärztlichen Leitlinie nicht nachlesen und fragt Sie, ob es nicht auch „etwas Verständliches“ für Patienten gebe. Da der Patient akute Schmerzen hat, können Sie ihn nicht unbehandelt nach Hause schicken, damit er entsprechende Patienteninformationen lesen kann. Schlagen Sie ihm vor, zunächst eine orale Schmerztherapie zu versuchen und besprechen mit ihm Maßnahmen, die er selbst ergreifen kann, um seine gesundheitliche Situation zu verbessern. Dies ist sehr beratungsintensiv, denn Sie wissen, der Patient ist gedanklich auf „seine Spritze“ festgelegt. Signalisieren Sie ihm, dass Sie seine Befürchtungen, ohne Injektion keine schnelle Linderung zu erfahren, verstanden haben. Er

klären Sie Ihrem Patienten nochmals die Risiken der Injektionstherapie gegenüber dem erwarteten Nutzen.

18.2.2.3 Treffen einer gemeinsam verantworteten Entscheidung

Sie haben Ihrem Patienten verständlich und möglichst ohne Verwendung von Fachbegriffen erklärt, worin Nutzen und Risiken der Behandlungsoptionen liegen. Der Patient sollte nun für sein spezielles gesundheitliches Problem auf dem gleichen Informationsstand sein wie Sie selbst. Sie können nun mit der Behandlung beginnen. Schlagen Sie dem Patienten vor, sich selbst weiter zu informieren und geben Sie ihm Hinweise, wo er dies tun kann und was er dabei beachten sollte. Entsprechende Quellen und Informationsmaterialien finden Sie in der Toolbox „Patienteninformationen“ als Kopiervorlagen im Anhang des Buches.

18.2.2.4 Empfehlung von Informationsquellen für Laien an Ihren Patienten

Sie haben Ihrem Patienten aus dem „Werkzeugkasten“ die Übersichten über verlässliche Patienteninformationen sowie über die Qualitätsbewertung von Patienteninformationen mitgegeben. Er kann nun selbst verschiedene Informationen zum Rückenschmerz recherchieren, so zum Beispiel in den Patientenleitlinien der Universität Witten-Herdecke (<http://www.patientenleitlinien.de/Rueckenschmerz/rueckenschmerz.html>) im Patienten-Informationsdienst der deutschen Ärzteschaft (<http://www.patienten-information.de>) nach Eingabe des Stichwortes „rückenschmerz“ (s. Abb. 18.3).

Der Patient kann selbst noch einmal in Ruhe die verschiedenen Quellen nachlesen. Dort findet er auch, dass im Falle von Spritzen mit mehr Nebenwirkungen bzw. Komplikationen zu rechnen ist und kann nun endgültig die Therapieempfehlung seines Arztes, der zur oralen Medikamentengabe riet, mit tragen.

Patenten-Information.de - Gesundheitsinfos finden - Mozilla Firefox

http://www.patenten-information.de/content/gesundheitsinfos/?

Patenten-Information.de

Qualitätsgeprüfte Behandlungsinformationen für Patienten und Laien
Ärztliches Zentrum für Qualität in der Medizin

gesundheitsinfos

Gesundheitsinfos finden Informationsqualität Kontakte Projekte Wir über uns

sie sind hier: startseite > gesundheitsinfos finden

Gesundheitsinfos finden

Suchhilfen
Zum ersten Mal hier?
Patientenschulungen
Literatur
Verlässliche englischsprachige Informationen
Selbsthilfe,
Patientenberatungsstellen
& Verbraucherzentralen

Suchergebnis im Katalog

Rückenschmerz in

Insgesamt: 8 Treffer

Behandlungsinformationen

5 Suchergebnisse in Behandlungsinformationen

Informationen zu Ursachen und Therapiemöglichkeiten bei Rückenschmerzen

Deutsches Grünes Kreuz e.V.

<http://www.forum-schmerz.de/web/schmerzcontent/de/rueckenschmerzen.htm>

Patientenbroschüre der Techniker Krankenkasse zum Thema: Rückenschmerzen

Techniker Krankenkasse unter wissenschaftlicher Beratung der Arzneimittelkommission der deutschen Ärzteschaft

<http://www.akdae.de/45/Kreuzschmerzen.pdf>

Schmerzen (allgemein)

Gmünder ErsatzKasse GEK mit der wissenschaftlichen Unterstützung und Beratung durch die Arzneimittelkommission der deutschen Ärzteschaft

<http://www.akdae.de/45/Schmerzen.pdf>

Check-In: Zu empfehlen als Hintergrundinformation

Discern: 😊

Check-In: Nicht zu empfehlen

Discern: 😊

DISCERN-Legende

4-5
Gute bis sehr gute Qualität
Information weist minimale Mängel auf, die Mehrzahl der Fragen wurde mit 4 und höher bewertet.

3
Mittlere Qualität
Information weist eventuell wichtige, aber nicht beträchtliche Mängel auf.

1-2
Schlechte Qualität
Information weist beträchtliche Mängel auf, die Mehrzahl der Fragen wurde mit 2 und schlechter bewertet

Warten auf www.patenten-information.de...

Abb. 18.3: Qualitätsbewertete Informationen zum Rückenschmerz im Patienten-Informationdienst der deutschen Ärzteschaft

18.2.3 Szenario 3

Eine Patientin hat bei einem Gentest, den sie bezüglich einer anderen gesundheitsbezogenen Frage durchführen ließ, erfahren, dass bei ihr eine Huntington-Mutation nachgewiesen wurde und es höchstwahrscheinlich zum Ausbruch der Chorea Huntington-Erkrankung kommen wird. Symptome hat sie derzeit noch nicht. Sie weiß lediglich, dass es sich um eine sehr schwere Erkrankung handelt. Völlig verzweifelt kommt sie in Ihre Sprechstunde (Sie sind Ihr Hausarzt und vertrauter Ansprechpartner) und bittet Sie, durch eine Behandlung den Ausbruch der Erkrankung zu verhindern.

Sie stehen vor der Aufgabe

- ▲ selbst Informationen über das Thema einzuholen,
- ▲ Ihre eigenen Grenzen anzuerkennen,

- ▲ Spezialisten für das entsprechende gesundheitliche Problem zu finden und
- ▲ die Patientin zu begleiten.

Wo liegt das Problem?

Die Patientin erwartet von Ihnen, dass Sie den Ausbruch der Erkrankung verhindern. Sie waren in der Vergangenheit immer ihr Ansprechpartner, wenn es um gesundheitliche Fragen und Probleme ging. Sie haben die Erkrankungen der Patientin stets heilen können. Jetzt müssen Sie diese Erwartung enttuschen.

Manchmal besteht die einzig vorhandene Evidenz darin, jemanden zu kennen, der sich mit dem Problem auskennt.

Muir A Gray,
anlässlich der 6. Jahrestagung
des DNEBM 2005 in Berlin

Mögliche Lösungswege

1. Identifikation von Spezialambulanzen und Zentren, die sich auf diese Erkrankung spezialisiert haben.
2. Unterstützung der Patientin bei der Suche nach ebenfalls Betroffenen.
3. Begleitung auch nach der Überweisung in ein spezielles Zentrum bzw. eine Spezialambulanz.

18.2.3.1 Identifikation von Spezialambulanzen und Zentren, die sich auf diese Erkrankung spezialisiert haben

Spezialambulanzen, Labore und andere Einrichtungen für seltene Erkrankungen sind in ORPHANET, der Europäischen Datenbank für seltene Erkrankungen und seltene Medikamente, gelistet (<http://www.orpha.net>). Darüber hinaus fördert das Bundesministerium für Bildung und Forschung in Deutschland Netzwerke für seltene Erkrankungen, bei denen man ebenfalls auf der Suche nach entsprechenden Experten fündig wird (<http://www.bmbf.de/de/1109.php>).

Da bei seltenen Erkrankungen die Studienlage sehr schlecht ist und demzufolge auch keine Leitlinien existieren, kommt es gerade hier darauf an, die Erfahrungen mit und das Wissen um die Erkrankung zu bündeln. Das tun international auch die European Organization for Rare Diseases (EURORDIS, <http://www.eurordis.org/>) und das Office of rare diseases des Nationale Institute of Health (<http://rarediseases.info.nih.gov>).

18.2.3.2 Unterstützung der Patientin bei der Suche nach anderen Betroffenen

Der Austausch mit ebenfalls Betroffenen kann besonders im Falle seltener Erkrankungen helfen, die Erkrankung zu verstehen und Therapieerfahrungen auszutauschen. Im vorliegenden Fall könnten Sie raten, sich an die Deutsche Huntington Hilfe e.V. (<http://www.huntington-hilfe.de>) zu wenden.

Anmerkung: Der Ansprechpartner in Deutschland für seltene Erkrankungen ist die

Allianz Chronischer Seltener Erkrankungen (ACHSE; <http://www.achse-online.de/>). Auch wenn es zur betreffenden Erkrankung noch keine Selbsthilfegruppe geben sollte (für Chorea Huntington ist dies der Fall), wird gibt es Unterstützung bei der Suche nach anderen Erkrankten. Inzwischen gibt es auch Selbsthilfegruppen, die sich virtuell austauschen (z.B. zur Asherman-Erkrankung).

18.2.3.3 Begleitung auch nach der Überweisung in ein spezielles Zentrum bzw. eine Spezialambulanz

Die Begleitung der Patientin auch nach der Überweisung in eine Spezialeinrichtung ist für Sie als Hausarzt mit Lotsenfunktion ohnehin eine wichtige Aufgabe, denn bei Ihnen laufen alle Informationen über die Krankengeschichte der Patientin zusammen. Im Kontakt mit den behandelnden Ärzten der Spezialeinrichtung können Sie begleitende Maßnahmen zur günstigen Beeinflussung des Allgemeinbefindens Ihrer Patientin durchführen.

18.3 Diskussion

Informationen sind das „Rückgrat“ im Prozess der Evidenzvermittlung an Patienten und stellen eine wichtige Säule im gemeinsamen Entscheidungsprozess von Arzt und Patient dar. Sie müssen jedoch stets im Kontext des individuell vorliegenden gesundheitlichen Problems und des Informationsbedürfnisses des jeweiligen Patienten gesehen werden.

Evidenzbasierte und patientenverständliche Informationen stehen im deutschsprachigen Raum bisher nur in einem begrenzten Umfang zur Verfügung. Ihre Qualität kann ad hoc vom medizinischen Laien nicht ohne weiteres beurteilt werden. Instrumente zur Bewertung der Informationsqualität untersuchen in erster Linie formale Aspekte und helfen somit, die Qualität und Sorgfalt

bei der Informationserstellung und -aufbereitung zu analysieren. Die Nutzer werden ermuntert, beim Fehlen zentraler Aussagen entsprechende zusätzliche Informationen einzuholen.

Ob eine Information tatsächlich auf der zum Zeitpunkt ihrer Erstellung bestverfügbaren Evidenz beruht, kann ein Instrument zur Qualitätsbewertung von Informationen allerdings nicht „messen“ [13]. Jedoch dienen international akzeptierte Qualitätskriterien für medizinische Information und Internetangebote als Informationsfilter. Qualitätssicherung von Gesundheitsinformationen wird dabei auf verschiedenen Ebenen praktiziert, die nicht alternativ, sondern durchaus additiv zu verstehen sind – Selbstverpflichtung der Anbieter zur Einhaltung definierter Qualitätskriterien bei der Erstellung und Verbreitung von Informationen (z.B. www.hon.ch, www.afgis.de), Kontrollen

und Zertifizierungen durch entsprechend autorisierte Einrichtungen (z.B. www.hon.ch, www.urac.org, www.medcircle.org, www.medicq.org, www.patienten-information.de, www.iqwig.de) und Qualitätsbewertungen durch die Nutzer selbst (z.B. www.discern.de, www.quick.org.uk). Unsere Studien zeigen, dass die Nutzer sehr schnell in der Lage sind, anhand des Instrumentes die Qualität von Materialien zu bewerten und dass sie schon nach kurzer Zeit sehr sicher in der Bewertung werden [14]. Unter der Perspektive, das Empowerment der Nutzer zu fördern und ihnen Möglichkeiten zu geben, unabhängig von Experten Qualitätsbewertungen vorzunehmen, ist die Dissemination von Checklisten wie DISCERN zu forcieren (s. Toolbox „Qualität von Patienteninformationen“ im Anhang des Buches).

Unter Berücksichtigung der Rolle der Evidenz in der Arzt-Patienten-Beziehung erge-

Tab. 18.2: Informationsaspekte

	Was soll unterstützt werden?	Voraussetzungen/Hilfen
Informationsaspekt	Kenntnis der aktuellen Datenlage zum betreffenden Problem	✂ Toolbox „Evidenzquellen für den Arzt“ Siehe im Anhang des Buches oder unter www.ebm-netzwerk.de/grundlagen/arbeitsmaterialien
	Kenntnis von Quellen verlässlicher Patienteninformationen	✂ Toolbox „Verlässliche Quellen für Patienten“ Siehe im Anhang des Buches oder unter www.ebm-netzwerk.de/grundlagen/arbeitsmaterialien
	Einschätzung der Zuverlässigkeit und Relevanz ärztlicher Evidenzquellen	S. Kap. 2
	Einschätzung der Zuverlässigkeit und Relevanz von Patienteninformationen als Hilfestellung für Patienten	✂ Toolbox „Qualität medizinischer Informationen für Patienten – Worauf ist zu achten?“ Siehe im Anhang des Buches oder unter www.ebm-netzwerk.de/grundlagen/arbeitsmaterialien
	Dokumentationshilfen zur Erhebung der Patientenerfahrung mit der Therapie (Schmerztagebücher, Tumornachsorgepass etc.)	✂ Toolbox „Dokumentationshilfen für Patienten“ Siehe im Anhang des Buches oder unter www.ebm-netzwerk.de/grundlagen/arbeitsmaterialien
	Empfehlung weiterführender Hilfen und Unterstützungsangebote	✂ Toolbox „Hinweis auf Patientenorganisationen und Selbsthilfe“ Siehe im Anhang des Buches oder unter www.ebm-netzwerk.de/grundlagen/arbeitsmaterialien

ben sich heute neue Erwartungen und Anforderungen an die ärztliche Kompetenz, die insbesondere die Bereiche der medizinischen, wissenschaftlich-methodischen sowie der psychosozialen und kommunikativen Fähigkeiten beinhalten [15]. Ärzte sind heute nicht nur gefordert, medizinisch wissenschaftliche Informationen für den eigenen Gebrauch zu recherchieren, sie müssen darüber hinaus auch dem Informationsbedürfnis und Informationsverhalten ihrer Patienten Rechnung tragen. Dies ist eng verknüpft mit den Erwartungen, die Patienten an die ärztliche Tätigkeit haben und kann im Falle einer Nichtbeachtung zu empfindlichen Kommunikationsstörungen bis hin zum Vertrauensverlust führen, die mit einem Arztwechsel verbunden sein können. Deshalb sollten Sie das aktive Informationsverhalten der Patienten entsprechend würdigen und unterstützen, indem Sie:

- ▲ Qualitätskriterien für medizinische Laieninformationen und Internetangebote kennen und an Ihre Patienten weitergeben,
- ▲ Verlässliche Quellen kennen und empfehlen können,

- ▲ Bewertungsinstrumente an die Patienten weitergeben,
- ▲ und mit Ihren Patienten gemeinsame Therapieentscheidungen treffen.

18.4 „Werkzeugkasten“ für Arzt und Patient

Was Sie bei der Information und Kommunikation mit Ihren Patienten noch beachten können und sollten, ist im Kapitel 5 unter der Überschrift „Welche Evidenz will der Patient?“ zusammengestellt. Darüber hinaus haben wir unter der Bezeichnung „Werkzeugkasten“ für Arzt und Patient eine Reihe von praktischen Hilfen, Informations- und Kopiervorlagen zusammengestellt, die Sie im Anhang des Buches finden. Diese Materialien stehen auch auf den Internetseiten des Deutschen Netzwerk Evidenzbasierte Medizin e.V. (DNEbM) unter der Adresse: www.ebm-netzwerk.de/grundlagen/arbeitsmaterialien zum Download zur Verfügung.

Tab. 18.3: Kommunikationsaspekte

	Was soll unterstützt werden?	Voraussetzungen/Hilfen
Kommunikationsaspekt	Vertrauen aufbauen durch gegenseitige Information (Patienten ermutigen, Fragen zu stellen; Gesprächsbereitschaft signalisieren; Abbau möglicher Ängste fördern etc.)	✂ Toolbox „Unterstützung des Arzt-Patienten-Gesprächs“ Siehe im Anhang des Buches oder unter www.ebm-netzwerk.de/grundlagen/arbeitsmaterialien
	Vertrauen aufbauen durch guten Kommunikationsstil	Partnerschaftlicher Umgang mit den Patienten, außer Patienten wünschen eher einen paternalistischen Stil (Siehe dazu auch die Typologie der Patienten in Kapitel 5 „Welche Evidenz will der Patient?“) Patienten sind auch „Fachleute“ ihrer ganz persönlichen Krankheitsgeschichte
	Vertrauen aufbauen durch ungestörte Atmosphäre	Respektierung der Privat- und Intimsphäre der Patienten (keine Störungen während der Untersuchung oder des Gesprächs durch Arzthelferinnen, Telefonate etc.; Sichtschutz beim An- und Auskleiden; Sichtschutz an den Fenstern etc.)

Literatur

- [1] Härter M, Partizipative Entscheidungsfindung (Shared Decision Making) – ein von Patienten, Ärzten und der Gesundheitspolitik geforderter Ansatz setzt sich durch. *Z Arztl Fortbild Qualitätssich.* (2004), 98, 89–92
- [2] Gemeinsamer Bundesausschuss (G-BA). Aufgaben des Gemeinsamen Bundesausschusses zur Qualitätssicherung im Gesundheitswesen. http://www.g-ba.de/cms/front_content.php?idcat=28 (14.07.05)
- [3] Coulter A, Magee E, (Ed.) (2003) *The European patient of the future.* Open Univ. Press, Maidenhead
- [4] Chang YC, Gesundheit in Fernsehen und Internet. Eine Studie über Angebot, Qualität und Nutzung von Gesundheitserziehung in den Medien. Dissertation zur Erlangung des Doktorgrades der sozialwissenschaftlichen Fakultät der Georg-August-Universität zu Göttingen. <http://webdoc.gwdg.de/diss/2001/chang/chang.pdf> (11.07.06)
- [5] Dierks ML et al. (2006) Bürger- und Patientenorientierung in Deutschland. Gesundheitsbericht für Deutschland. Robert Koch Institut, Berlin
- [6] Schaeffer D, Dierks ML (2006) Patientenberatung in Deutschland. In: Schaeffer D, Ewers M, Krause H, Patientenberatung, 71–96. Huber Verlag, Bern
- [7] Dierks ML, Kurtz V, Seidel G (2006) Versorgungsforschung im Bereich der unabhängigen Verbraucher- und Patientenberatung. Evaluation der Nutzerstrukturen und der Beratungsqualität. In: Hey M, Maschewsky-Schneider U, Kursbuch Versorgungsforschung, 192–205. Med.Wiss.Verl.-gesellsch., Berlin
- [8] Dierks ML, Diel F, Schwartz FW (2003) Stärkung der Patientenkompetenz: Information und Beratung. In: Schwartz FW et al., *Das Public Health Buch*, 229–348. Urban & Fischer, München
- [9] Koch O, Kaltenborn R, Wissensmanagement am Arbeitsplatz. Mehr Zeit für Patienten durch bessere Information. Dt Arztebl online. (2005)
- [10] Sänger S et al., Manual Patienteninformatio. Empfehlungen zur Erstellung evidenzbasierter Patienteninformatio. http://www.patienten-information.de/content/download/manual_patienteninformatio_04_06.pdf (23.05.06)
- [11] Charnock D et al., DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health.* (1999), 53, 105–111
- [12] Sänger S et al., Gut informiert über Gesundheitsfragen, aber wie? Das Deutsche Clearingverfahren für Patienteninformatio – Zielsetzung, Hintergrund und Arbeitsweise. *Gesundheitswesen.* (2002), 64, 391–397
- [13] Steckelberg A et al., Kriterien für evidenzbasierte Patienteninformatio. *Z Arztl Fortbild Qualitätssich.* (2005), 99, 343–351
- [14] Dierks ML et al. (2001) Medizinische Informationsangebote im Internet. Bewertung durch Experten und Laien im Vergleich. Abschlussbericht.
- [15] Ollenschläger G, Die Kunst, heutzutage ein guter Arzt zu sein. *Ärztebl Bad-Würt.* (2004), 59, 111–114

19 Von der Evidenz zur Empfehlung

Regina Kunz, Monika Lelgemann, Gordon Guyatt, Gerd Antes, Yngve Falck-Ytter, Holger Schünemann

Leitlinien sind ein Kernelement für den Wissenstransfer im Gesundheitswesen und daher aus der Versorgungsrealität nicht mehr wegzudenken. Die Leitliniengruppen der Fachgesellschaften und anderer Organisationen haben Kompetenzen in der Methodik der Leitlinienerstellung erworben, um sicherzustellen, dass die empfohlenen Vorgehensweisen den Patienten mehr nutzen als schaden [1]. Zu diesen Kernkompetenzen gehört, die richtigen Fragen zu stellen und die beste verfügbare Evidenz aufzusuchen, diese auf Glaubwürdigkeit und klinische Relevanz zu überprüfen und die Übertragbarkeit auf die Versorgungssituation zu reflektieren, um schließlich einen Konsens für die Empfehlungen an Ärzte und Patienten zu finden.

Das Vorgehen entspricht dem internationaler Leitliniengruppen. Wie diese, sehen sich auch die deutschen Leitliniengruppen mit unzureichend geklärten Fragen, Problemen und Hürden der praktischen Umsetzung der vereinbarten Vorgehensweisen konfrontiert: Welche Konsequenzen hat die differenzierte Qualitätsbewertung für die Entscheidungsfindung, wenn die Studien, wie häufig der Fall, nicht richtig gut, aber auch nicht ganz offensichtlich schlecht sind? Wie kommt man von der Evidenz zur Empfehlung? Welche Faktoren neben der Evidenz beeinflussen die Empfehlung, und lassen sich diese nachvollziehbar und reproduzierbar operationalisieren? Daran knüpft sich zu guter Letzt die Frage, wie sich eine Empfehlung rasch und unmissverständlich an Ärzte und Patienten kommunizieren lässt.

Um Fragen dieser Art zu lösen, hat eine unabhängige internationale Gruppe aus Leit-

linienerstellern, Methodikern und Klinikern – die GRADE Working Group – die bestehenden Methodiken zur Leitlinienerstellung weiterentwickelt und ein System erarbeitet, das anhand expliziter Kriterien eine transparente Bewertung der Evidenz und eine transparente Umsetzung der Studienevidenz in differenzierte Empfehlungen ermöglicht. GRADE Working Group steht dabei für Grades of Recommendations, Assessment, Development and Evaluation. 2004 wurde das System erstmals der Öffentlichkeit vorgestellt und seither weiter entwickelt und modifiziert [2–4]. Die rasche Akzeptanz des GRADE-Systems von bedeutenden medizinischen Gesellschaften mit langjähriger Erfahrung in Leitlinienerstellung wie dem American College of Chest Physicians [3] (die 1986 als erste wissenschaftliche Gesellschaft weltweit ihren Guidelines Evidenzstufen zugrunde gelegt haben [5]), der Weltgesundheitsorganisation WHO [6], der American Thoracic Society [2]; der KDIGO-Initiative (Kidney Disease: Improving Global Outcomes Initiative) [7]; dem American College of Physicians, der American Endocrine Society und dem elektronischen Lehrbuch „UpToDate“ [8] (www.uptodate.org) belegen den Mehrwert der GRADE-Methodik.

Da bereits mehrere exzellente Beschreibungen der GRADE-Methodik existieren, haben wir uns bei der Konzeption des vorliegenden Kapitels entschieden, aus den Schlüsselpublikationen [2–4] eine modifizierte deutsche Übersetzung zu erstellen. In einem weiteren Abschnitt gehen wir gezielt auf häufige Fragen deutscher Leitliniengruppen ein, die sich möglicherweise mit der

GRADE-Methodik lösen lassen. Aus Platzgründen können wir nicht alle Aspekte der Originalpublikationen abdecken und haben deshalb auf der Homepage der GRADE Working Group (www.gradeworkinggroup.org) eine Vielzahl von Texten, zusätzlichen Tabellen und vor allem Beispiele eingestellt, die das theoretische Konzept weiter veranschaulichen. In unserem Buchkapitel haben wir die entsprechenden Stellen, zu denen weiterführende Informationen auf der Webseite vorliegen, gekennzeichnet.

19.1 Die GRADE-Methodik: Mehr Transparenz bei der Erstellung von abgestuften Handlungsempfehlungen

Empfehlungen für die Patientenbehandlung bedeuten immer das Abwägen zwischen dem möglichen Nutzen einer Maßnahme und ihren Nachteilen, d.h. möglichem Schaden, Belastungen und Kosten für den Patienten oder seine Angehörigen. Um Leitlinien, z.B. einer Fachgesellschaft, angemessen in den Kontext ihres eigenen Patienten stellen zu können, müssen Ärzte die zugrunde liegende Evidenz und die Bewertungen verstehen, die in die Formulierung der Empfehlungen eingegangen sind. Ein transparenter systematischer Ansatz bei der Entwicklung von Handlungsempfehlungen kann helfen, Bias zu minimieren und die Interpretation der Evidenz zu erleichtern. Dieses Kapitel bietet einen Leitfaden, wie Leitliniengruppen mit dem GRADE-System die Qualität der Evidenz bewerten und durch differenzierte standardisierte Formulierungen den Stärkegrad ihrer Empfehlungen kommunizieren können und welche Konsequenzen sich daraus für Ärzte und Patienten für die Behandlung ergeben.

Ausgangspunkt ist ein systematischer Review oder eine Meta-Analyse. Mit dem System der GRADE Working Group lässt sich die darin eingeschlossene Evidenz nach ihrer

Qualität klassifizieren und man kann je nach Stärke der Aussage abgestufte Empfehlungen entwickeln [9; 10]. Dabei hat sich die GRADE Working Group unter anderem der Herausforderung gestellt, für die Bewertung der Evidenz und der resultierenden Formulierung der Empfehlungen explizite und transparente Kriterien zu entwickeln.

Bei der Entwicklung von Leitlinien beginnt das GRADE-System mit einer Überprüfung der Qualität der Evidenz, an die sich Schlussfolgerungen über die Stärke der Empfehlung anschließen. Da Ärzte in erster Linie auf Empfehlungen zurückgreifen, um zu wissen, welche diagnostische oder therapeutische Vorgehensweise für ihren Patienten am besten geeignet ist, formuliert das GRADE-System zuerst eine Empfehlung mit der zugehörigen Empfehlungsstärke, gefolgt von einer Aussage zur Qualität der Evidenz. Die bewusste Trennung von Überprüfung der Studienqualität einerseits und der Vergabe von Empfehlungsstärken andererseits ist ein essenzielles und definierendes Merkmal dieses Graduierungssystems.

19.1.1 Die Stärke einer Empfehlung

Wenn Leitliniengruppen mit GRADE eine Empfehlungsstärke festlegen, treffen sie eine Aussage, wo ihrer Überzeugung nach eine Behandlung in dem Spannungsfeld zwischen erwünschten Effekten und unerwünschten Auswirkungen anzusiedeln ist und wie sicher sie sind, dass die Einschätzung wissenschaftlich fundiert ist. Erwünschte Effekte können sich dabei auf die angestrebte Wiederherstellung der Gesundheit („Health Outcomes“), auf eine geringere Belastung durch die Behandlung für den Patienten oder seine Angehörigen, oder auf einen verminderten Einsatz von Ressourcen (im Allgemeinen Kosten) beziehen. Unerwünschte Effekte können seltene schwere Nebenwirkungen, andere unerwünschte

Wirkungen, größere Belastungen für Patienten oder Angehörige und einen höheren Verbrauch an Ressourcen bedeuten. Patienten können z.B. Empfehlungen als eine Belastung empfinden, wenn sie Maßnahmen umfassen, die der Patient oder seine Familie eher vermeiden wollen, wie die Einnahme von Medikamenten, Verhaltensänderungen, lästige regelmäßige Laboruntersuchungen oder Arztbesuche. Zeigt sich, dass die erwünschten Effekte einer Intervention ihre unerwünschten Auswirkungen überwiegen, werden Leitliniengruppen empfehlen, diese Maßnahme den „typischen“ Patienten anzubieten.

19.1.2 Starke und schwache Empfehlungen

Sind sich Leitliniengruppen sicher, dass aus der Befolgung einer Empfehlung erwünschte Effekte resultieren und diese die unerwünschten Effekte überwiegen (oder umgekehrt), formulieren sie eine starke Empfehlung (s. Abb. 19.1). Eine starke Empfehlung erfordert gewöhnlich hochwertige Evidenz mit einer präzisen Effektschätzung der Vor- und der Nachteile, sowie eine eindeutige Stellungnahme zugunsten (oder zuungunsten) des Nutzens verglichen mit den Nachteilen einer Intervention. Leitliniengruppen formulieren in der Regel abgeschwächte

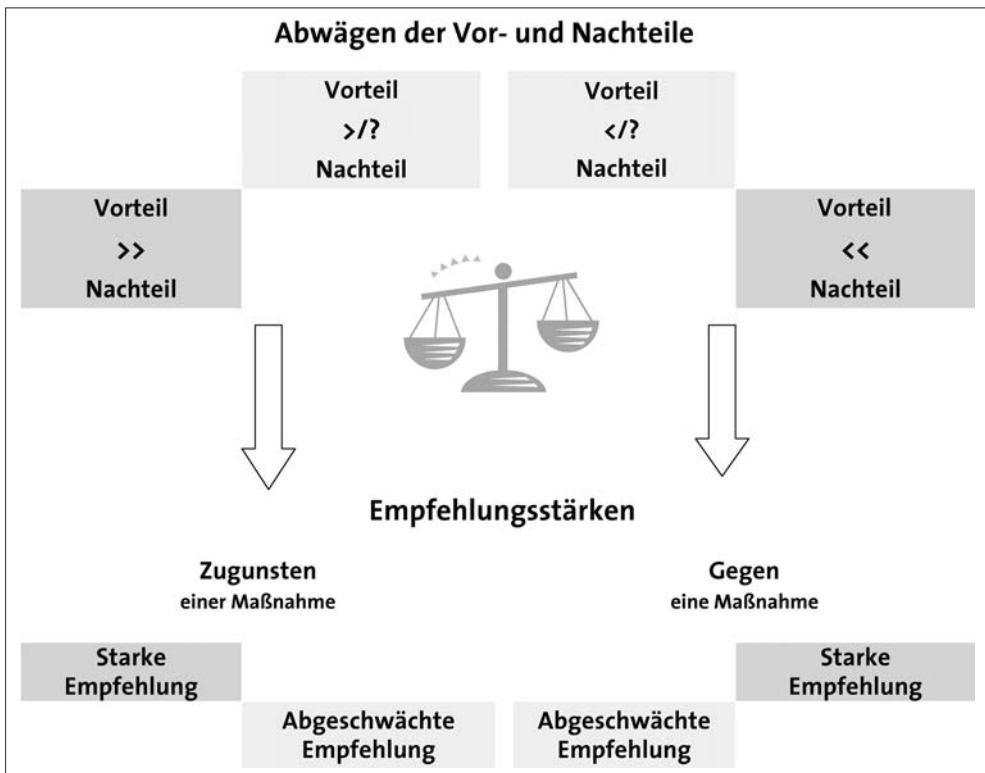


Abb. 19.1: Empfehlungsstärken. Die Abbildung beschreibt die Abwägung zwischen erwünschten und unerwünschten Effekten, die einer Empfehlung zugrunde liegen. Resultiert aus der Abwägung eine eindeutige Aussage, so folgt eine starke Empfehlung zugunsten oder zuungunsten einer Maßnahme (links und rechts außen). Ist das Ergebnis der Abwägung nicht eindeutig, so folgt eine schwache Empfehlung zugunsten oder zuungunsten einer Maßnahme (Mitte). Ein breites Spektrum an vorherrschenden Wertvorstellungen kann die Balance zwischen erwünschten und unerwünschten Effekten deutlich ändern und zu weniger eindeutigen Aussagen führen.

Empfehlungen, wenn eine nennenswerte Unsicherheit über die Größe der Vorteile und/oder des potenziellen Schadens besteht (in der Regel als Folge von geringwertiger Evidenz) oder Vorteile und potenzieller Schaden sehr eng beieinander liegen.

Ärzte, die sich bemühen, Entscheidungen gemeinsam mit ihren Patienten zu treffen, werden sich zunehmend über die Bedeutung von Patientenwerten und Patientenpräferenzen bewusst. Die Perspektive des Patienten, seine Werte und Vorstellungen können ein weiterer Ausgangspunkt sein, um Empfehlungen in einer starken oder abgeschwächten Form zu formulieren (s. Tab. 19.1). Sind die Vorteile einer Behandlung so offensichtlich (oder im umgekehrten Fall, die Nachteile den Vorteilen weit überlegen), dass sich nahezu alle Patienten für das gleiche Vorgehen entscheiden, können Leitliniengruppen eine starke Empfehlung aussprechen [s. GRADE-Homepage, Beispiel 1: Myokardinfarkt und Aspirin].

Charakteristisch für eine „starke Empfehlung“ ist, dass der Behandlungsvorteil so eindeutig und offensichtlich ist, dass Ärzte ihren Patienten (dem „typischen“ Patienten) die Behandlung einfach erklären können mit dem Hinweis, dass sie von der Behandlung

profitieren werden. Bei abgeschwächten Empfehlungen müssen Ärzte dagegen wesentlich sorgfältiger Nutzen, Schaden und Belastungen gemeinsam mit ihrem Patienten abwägen und die Entscheidung für oder gegen eine Behandlung mit dessen persönlichen Sichtweisen und Wertvorstellungen in Einklang bringen. Solche Situationen treten typischerweise dann auf, wenn eine nennenswerte Anzahl von Patienten aufgrund von unterschiedlichen Wertvorstellungen trotz gleicher Evidenzlage zu unterschiedlichen Entscheidungen kommt [s. GRADE-Homepage, Beispiel 2: Tiefe Beinvenen-Thrombose und Marcumar/Aspirin].

19.1.3 Welche Faktoren wirken sich auf die Stärke einer Empfehlung aus?

Tabelle 19.2 beschreibt die Faktoren, die Leitliniengruppen bei ihrer Entscheidung über die Stärke einer Empfehlung berücksichtigen sollen. Die angesprochenen Kriterien lassen sich auf jede der drei Säulen: Nutzen, Schaden und Belastung der Therapie anwenden.

Das erste Kriterium „Methodische Qualität der Evidenz“ in Tabelle 19.2 ist sehr umfangreich und wir haben diesem Punkt des-

Tab. 19.1: Wie lassen sich starke oder abgeschwächte Empfehlungen verstehen?

Starke Empfehlungen zugunsten einer Maßnahme bedeuten ...

für Patienten: Die meisten Patienten würden die vorgeschlagene Vorgehensweise befolgen, und nur wenige würden sie ablehnen.

Formale Entscheidungshilfen wie „Decision-Aids“ werden dem Patienten bei der Entscheidungsfindung kaum weiterhelfen.

für Ärzte: Die meisten Patienten sollen die Intervention erhalten.

Abgeschwächte Empfehlungen zugunsten einer Maßnahme bedeuten ...

für Patienten: Die Mehrheit der Patienten würde in dieser Situation die vorgeschlagene Vorgehensweise befürworten, aber viele würden sich auch dagegen entscheiden.

für Ärzte: Wahrscheinlich können formale Entscheidungshilfen wie Decision-Aids dem Patienten weiterhelfen, sich über seine eigenen Wertvorstellungen klar zu werden und diese in seine Entscheidung zu integrieren.

Ärzte sollten die erwünschten und unerwünschten Effekte der Intervention im Detail kennen und diese mit den (Wert-)Vorstellungen ihres Patienten in Einklang bringen.

Tab. 19.2: Faktoren, die die Stärke einer Empfehlung beeinflussen

Kriterium	Empfohlener Prozess und Beispiele
A. Qualität der Evidenz	
1. Qualität der Evidenz	<p>Starke Empfehlungen erfordern prinzipiell hochwertige Qualität der Evidenz für alle entscheidenden Gesundheitszustände (Outcomes). Mit abnehmender Qualität sinkt die Wahrscheinlichkeit, dass aus der Evidenz eine starke Empfehlung resultiert.</p> <p>Beispiel:</p> <p>Viele hochwertige randomisierte Studien belegen den Nutzen von inhalierten Steroiden bei Asthma. Dagegen ist der Nutzen der Pleurodesse bei Pneumothorax nur in Fallserien belegt.</p>
B. Abwägung von erwünschten und unerwünschten Effekten	
2. Relative Bedeutung der Gesundheitszustände (Outcomes)	<p>Leitliniengruppen sollten dokumentieren, welche relative Bedeutung sie den einzelnen patientenrelevanten Endpunkten in der Leitlinie zuordnen. GRADE kennt Outcomes die entscheidungstragend sind, Outcomes, die klinisch wichtig, aber nicht entscheidungstragend sind und Outcomes, die aus klinischer Sicht unbedeutend sind. Messen Patienten den Outcomes sehr unterschiedliche Bedeutung bei, werden starke Empfehlungen weniger wahrscheinlich.</p> <p>Beispiel:</p> <p>Bei einem Vergleich der kombinierten Strahlen-/Chemotherapie mit alleiniger Strahlentherapie bei Patienten mit nicht resezierbarem lokal fortgeschrittenen nicht-kleinzelligen Lungenkarzinom führt die kombinierte Behandlung zu einer Lebensverlängerung von einigen Monaten, aber um den Preis einer erheblichen Toxizität durch die Chemotherapie und weiterer Belastungen für den Patienten. Es wird Patienten geben, die einer Lebensverlängerung von einigen Monaten einen hohen Wert beimessen und dafür erhebliche Belastungen in Kauf nehmen, während andere Patienten möglicherweise weniger bereit sind, die schweren Nebenwirkungen und Beeinträchtigungen für die geringe Lebensverlängerung zu ertragen.</p>
3. Grundrisiko von Patientengruppen, ein unerwünschtes Ereignis zu erleiden	<p>Leitliniengruppen sollen berücksichtigen, welches Risiko Patienten haben, ohne Behandlung ein unerwünschtes Ereignis zu erleiden („Grundrisiko“). Falls sich das Grundrisiko für ein Ereignis in verschiedenen Patientenpopulationen nennenswert unterscheidet, sollten die jeweiligen Populationen getrennte Empfehlungen erhalten. Je höher das Grundrisiko und je größer der Behandlungseffekt, desto größer die Wahrscheinlichkeit für eine starke Empfehlung.</p> <p>Beispiel:</p> <p>Einige chirurgische Patienten haben ein sehr geringes Risiko für eine postoperative tiefe Beinvenenthrombose/Lungenembolie. Hier erscheint die Thrombembolie-Prophylaxe fragwürdig oder unnötig, während andere chirurgische Patienten ein wesentlich höheres Risiko für Thrombosen/Embolien haben, was eine Thrombembolie-Prophylaxe unbedingt erfordert.</p>
4. Größe des Relativen Risikos	<p>Große <u>relative</u> Effekte führen mit größerer Wahrscheinlichkeit zu einer starken Empfehlung zugunsten der Behandlung und umgekehrt.</p> <p>Beispiel:</p> <p>Im Vergleich zu Aspirin führt Clopidogrel nach einer TIA nur zu einer geringen zusätzlichen Absenkung des Schlaganfallrisikos (Relative Risiko-Reduktion (RRR) 8,7%), während sich z.B. das Schlaganfallrisiko bei Patienten mit Vorhofflimmern durch Marcumar verglichen mit Placebo um 68% (RRR) reduzieren lässt.</p>

Tab. 19.2: Fortsetzung

Kriterium	Empfohlener Prozess und Beispiele
5. Absolute Größe des Effekts	Große <u>absolute</u> Behandlungseffekte führen mit größerer Wahrscheinlichkeit zu starken Empfehlungen zugunsten der Behandlung Beispiel: Bei Patienten mit einem jährlichen Schlaganfall-Risiko von 4% führt die Prophylaxe mit Clopidogrel im Vergleich zu Aspirin zu einer absoluten Schlaganfall-Reduktion von 0,3% pro Jahr. Patienten mit Vorhofflimmern und vergleichbarem Schlaganfallrisiko erfahren durch Marcumar (statt Aspirin) eine absolute Schlaganfall-Reduktion von 2% pro Jahr.
6. Präzision der Effektschätzung	Je präziser der Behandlungseffekt (d.h. je enger das Konfidenzintervall), desto besser die Voraussetzungen für eine starke Empfehlung. Beispiel: Aspirin vs. Placebo zur Schlaganfallprophylaxe bei Patienten mit Vorhofflimmern weist ein breiteres Konfidenzintervall auf als Aspirin vs. Placebo bei Patienten mit TIA.
7. Kosten (optional)	Ein bedeutender Effekt sollte immer zu vertretbaren Kosten erreicht werden. Je höher die Mehrkosten einer alternativen Behandlungsform bei sonst gleichen Faktoren, desto weniger wahrscheinlich wird die Entscheidung für eine starke Empfehlung zugunsten der Intervention ausfallen. Beispiel: Bei Patienten mit TIA führt Clopidogrel verglichen mit Aspirin zu wesentlich höheren Kosten.

halb unter 19.1.5 einen eigenen Absatz gewidmet. Das zweite Kriterium „Relative Bedeutung der Endpunkte“ besagt, dass Leitliniengruppen bei Endpunkten von hoher Patientenrelevanz („entscheidende Endpunkte“) im Allgemeinen stärkere Empfehlungen verabschieden sollen, als wenn es um Endpunkte geht, die aus Patientensicht weniger bedeutsam sind.

Mit dem Beispiel von dosisangepasstem Marcumar verglichen mit Aspirin zur Prävention von Schlaganfall bei Patienten mit Vorhofflimmern lassen sich gleich mehrere Faktoren demonstrieren, die die Stärke einer Empfehlung beeinflussen. Eine Meta-Analyse zeigte bei jeder Art von Schlaganfall eine Relative Risiko-Reduktion (RRR) von 46% durch Marcumar verglichen mit Aspirin [11]. Auf diesen großen Effekt stützt sich eine starke Empfehlung für die Behandlung mit Marcumar. Das relativ enge 95%-Konfidenzintervall (RRR 29%–57%) belegt, dass Marcumar

im ungünstigsten Fall mindestens zu einer RRR von 29% führt, was die starke Empfehlung zusätzlich untermauert. Gleichzeitig bringt Marcumar jedoch auch erhebliche Belastungen für den Patienten: die Zufuhr von Vitamin K in der täglichen Ernährung muss konstant gehalten werden, die Intensität der Antikoagulation muss mit regelmäßigen Bluttests überwacht werden, und der Patient lebt in dem Bewusstsein eines erhöhten Risikos für kleinere und größere Blutungen. Allerdings fürchten die meisten Patienten einen Schlaganfall wesentlich stärker als eine Blutung [12]. Aus diesem Grund würden nahezu alle Patienten mit hohem Schlaganfallrisiko Marcumar wählen. In dieser Situation wäre eine starke Empfehlung angemessen.

Das Grundrisiko eines Patienten für ein negatives Ereignis, das durch die Behandlung gesenkt werden soll, kann für die Formulierung von Empfehlungsstärken ebenfalls elementar sein. Nehmen Sie eine 65-jäh-

rige Patientin mit Vorhofflimmern ohne sonstige Risikofaktoren für Schlaganfall. Ihr Grundrisiko für einen Schlaganfall im nächsten Jahr liegt bei ungefähr 2%. Verglichen mit Aspirin kann dosisangepasstes Marcumar das Risiko auf ungefähr 1% senken. Die meisten Patienten mit ausgeprägter Furcht vor einem Schlaganfall werden finden, dass die Vorteile einer Marcumar-Behandlung ihre Nachteile aufwiegen. Es wird jedoch Patienten geben, die die Vorteile nicht als ausreichend hoch erachten, um die Nebenwirkungen und die Unannehmlichkeiten der Therapie zu akzeptieren. Wenn voll informierte Patienten aufgrund unterschiedlicher individueller Präferenzen und Werte zu sehr unterschiedlichen Entscheidungen gelangen, sollten Leitliniengruppen abgeschwächte Empfehlungen abgeben.

Idealerweise können Ärzte die (Wert-)Vorstellungen direkt von ihren Patienten erheben, während Leitliniengruppen auf entsprechende populationsbezogene Studien zurückgreifen können. Allerdings sind diese Studien oft nicht verfügbar. In solchen Situationen können die Gruppen zu ihren Empfehlungen explizit dokumentieren, welche Werte und Vorstellungen sie den Empfehlungen zugrunde gelegt haben [s. GRADE-Homepage, Beispiel 3: Kombinationsbehandlung Chemotherapie plus Strahlentherapie versus alleinige Strahlentherapie].

Je enger erwünschte und unerwünschte Effekte einer Maßnahme beieinander liegen oder je größer die Unsicherheit des Effekts ist, desto stärker fallen bei der Entscheidung

über den Einsatz einer Behandlung die Kosten ins Gewicht. Die Frage der Ressourcen-Allokation in Abhängigkeit von der Evidenzlage ist jedoch ein eigenes Thema [13], auf das wir in diesem Kapitel nicht weiter eingehen wollen.

Beim Ausformulieren der Empfehlungen schlägt das GRADE-System vor, dass sich die Empfehlungsstärke auch in den Formulierungen der Leitliniengruppe widerspiegeln sollte. Dies lässt sich bei starken Empfehlungen mit den Begriffen „soll“ oder „wir empfehlen“ und bei schwächeren Empfehlungen mit dem Begriff „sollte“ oder „kann“ erzielen (s. auch Kap. 35).

19.1.4 Rolle der methodischen Qualität bei der Formulierung von Empfehlungsstärken

Im GRADE-System stammt Evidenz mit hoher Qualität von einer oder mehreren randomisierten kontrollierten Studien (RCTs), die, gut geplant und durchgeführt, konsistente und direkt anwendbare Ergebnisse geliefert haben. Hochwertige Evidenz kann außerdem von gut gemachten Beobachtungsstudien mit besonders großen Effekten kommen (s. Tab. 19.3). RCTs mit substanziellen Schwächen und gut gemachte Beobachtungsstudien mit großen Effekten reflektieren Evidenz von mittlerer Qualität. Gut gemachte Beobachtungsstudien mit kleinen Effekten und gelegentlich auch RCTs mit sehr schweren Fehlern werden als Evidenz von

Tab. 19.3: Studiendesigns und Qualitätsstufen

Zugrundeliegendes Studiendesign	Qualitätsstufe
Randomisierte kontrollierte Studien (RCTs)	Hoch
Herabgestufte RCTs oder hochgestufte Beobachtungsstudien	Mittel
Gut gemachte Beobachtungsstudien mit Kontrollgruppe & RCTs mit schwerwiegenden Fehlern	Niedrig
Andere Studien und Evidenzquellen (z.B. Fallberichte oder Fallserien)	Sehr niedrig

niedriger Qualität eingestuft. Beobachtungsstudien mit unzureichenden Kontrollgruppen und unsystematische klinische Beobachtungen (z.B. Fallserien oder Fallberichte) bilden die Gruppe mit sehr niedriger Qualität. Im weiteren wird das Kapitel ausführen, wie man Evidenz nach ihrer methodischen Qualität klassifiziert.

19.1.5 Welche Faktoren stufen die Qualität der Evidenz herab?

Die folgenden Schwächen beeinträchtigen die Qualität der Evidenz, die einer Empfehlung zugrunde liegt (s. Tab. 19.4).

▲ **Schwächen in der Methodik:** Weisen die Studien in ihrer Methodik deutliche Schwächen auf, die mit hoher Wahrscheinlichkeit zu Verzerrungen (Bias) in der Schätzung des Behandlungseffekts führen, sinkt das Vertrauen in die Angemessenheit einer Empfehlung. Im Kontext von RCTs handelt es sich dabei um Schwächen wie eine fehlende Intention-to-treat-Analyse, fehlende Verblindung bei subjektiven Endpunkten, ein hoher Verlust in der Nachbeobachtung oder RCTs, die man wegen eines sich abzeich-

nenden Nutzens vorzeitig abgebrochen hatte [s. GRADE-Homepage, Beispiel 4: Danaparoid bei Heparin-induzierter Thrombozytopenie (HIT)].

▲ **Unerklärte Heterogenität der Ergebnisse:** Wenn Studien zur gleichen Frage sehr unterschiedliche Behandlungseffekte liefern (Heterogenität oder Variabilität der Ergebnisse), sollten Forscher nach Ursachen suchen, die diese Heterogenität erklären. So könnten z.B. Medikamente bei kränkeren Patienten oder in höheren Dosen einen größeren relativen Effekt zeigen. Heterogenität ohne plausible Ursachen führt zu einem Abstufen der Evidenz [s. GRADE-Homepage, Beispiel 5: Pentoxifyllin bei Claudicatio Intermitens].

▲ **Indirekte Evidenz** (d.h. die Fragestellung der Leitliniengruppe und die vorhandene Evidenz unterscheiden sich deutlich in Bezug auf Population, Intervention und Vergleichsintervention sowie Endpunkt): Es existieren nur Studien mit ähnlichen, aber nicht identischen Populationen, Interventionen, Endpunkten o.Ä. wie die, für die Empfehlungen formuliert werden sollen [s. GRADE-Homepage, Beispiel 6: Vogelgrippe].

Tab. 19.4: Faktoren, die die Qualität der Evidenz abwerten

RCTs mit einer hohen Wahrscheinlichkeit von Bias aufgrund von unzureichender Studienplanung und -durchführung:
<ul style="list-style-type: none"> • schwerwiegende Einschränkungen → Abwertung um 1 Qualitätsstufe; • sehr schwerwiegende Einschränkungen → Abwertung um 2 Qualitätsstufen
Inkonsistente Ergebnisse, einschließlich Problemen in Untergruppen-Analysen → Abwertung um 1 Qualitätsstufe
Indirekte Evidenz, d.h. indirekte Vergleiche bzgl. der Populationen, Interventionen, Vergleichsintervention und der Outcomes
<ul style="list-style-type: none"> • und infolgedessen eine nennenswerte Ungewissheit bzgl. des tatsächlichen Effekts → Abwertung um 1 Qualitätsstufe; • und infolgedessen eine erhebliche Ungewissheit bzgl. des tatsächlichen Effekts → Abwertung um 2 Qualitätsstufen
Ungenauigkeiten in den Ergebnissen, d.h. weite Konfidenzintervalle → Abwertung um 1 Qualitätsstufe
Hohe Wahrscheinlichkeit von Reporting-Bias → Abwertung um 1 Qualitätsstufe

- ▲ **Fehlende Präzision:** Wenn Studien aufgrund von geringen Patientenzahlen und/oder wenigen Ereignissen weite Konfidenzintervalle aufweisen, wird eine Leitliniengruppe die Qualität der Evidenz wegen der resultierenden Unsicherheit über den tatsächlichen Effekt niedriger einstufen [s. GRADE-Homepage, Beispiel 7: Nadroparin bei Sinusvenenthrombose].
- ▲ **Reporting-Bias:** Die Qualität der Evidenz kann abgeschwächt werden, wenn Forscher bestimmte Studien überhaupt nicht publizieren (typischerweise Studien mit negativem Ergebnis) oder bestimmte Endpunkte nicht berichten (typischerweise bei einer möglicherweise schädigenden oder ganz fehlenden Wirkung). Leitliniengruppen können über die Existenz von Reporting-Bias häufig nur Mutmaßungen anstellen und haben dadurch nur eine begrenzte Handhabe, mit diesen Situationen umzugehen. Eine typische Konstellation für Reporting-Bias ergibt sich, wenn die publizierte Studienlage im Wesentlichen kleinere Studien umfasst, die alle durch die Industrie finanziert wurden [14]. So zeigten 14 Studien über Flavinoide bei Patienten mit Hämorrhoiden sehr große Effekte, aber sie rekrutierten insgesamt nur 1432 Patienten [15]. Wegen der erheblichen Rolle der Sponsoren in den meisten Studien stellt sich die Frage nach der Existenz von unpublizierten Studien, in denen kein Nutzen nachgewiesen wurde.
- ▲ Studien können mehr als nur eine Schwäche aufweisen. Je ausgeprägter die Mängel, desto geringer ist die Qualität der Evidenz. Existieren zu einer Frage z.B. nur randomisierte kontrollierte Studien mit gravierenden Schwächen, so transportieren diese nur Evidenz von sehr niedriger Qualität.

19.1.6 Welche Faktoren führen zu einer Heraufstufung der Qualität der Evidenz bei Beobachtungsstudien?

Auch Beobachtungsstudien können Evidenz von mittlerer oder hoher Qualität liefern. Gut gemachte Beobachtungsstudien werden prinzipiell als Evidenz von niedriger Qualität eingestuft. Dennoch gibt es bestimmte Umstände, unter denen Leitliniengruppen diese Studien als Evidenz von mittlerer Qualität oder gar hoher Qualität einstufen können (s. Tab. 19.5).

- ▲ Wenn methodisch hochwertige Beobachtungsstudien große oder sehr große Effekte nachweisen und sich diese großen Effekte konsistent in mehreren Studien finden, stärkt dies unser Vertrauen in die Ergebnisse. Obwohl Beobachtungsstudien den tatsächlichen Effekt häufig überschätzen, kann das schwache Studiendesign allein nicht den gesamten beobachteten Effekt erklären. Trotz aller Vorbehalte gegenüber Beobachtungsstudien können wir daher in gewisser Weise davon ausgehen, dass ein Effekt wirklich existiert. Tabelle 19.5 zeigt, wie sich Studien von der Stufe „niedrige Qualität“ auf die Stufe „mittlere oder hohe Qualität“ verschieben lassen, aber – und dies sei nochmals explizit vermerkt – nur unter der Bedingung, dass der Effekt groß ist und keine anderen methodischen Schwächen vorliegen. Diese Konstellation ist eher die Ausnahme, wie z.B. bei einer Meta-Analyse aus 37 Beobachtungsstudien über die Auswirkung der Embolie-Prophylaxe bei Herzklappenersatz, wo sich für Marcumar im Vergleich zu Aspirin ein relatives Risiko für Thrombembolie von 0.17 (95%-CI: 0.13–0.24) fand. Dieser überaus große Effekt erlaubt eine Zuordnung als hochwertige Evidenz [16].
- ▲ Gelegentlich unterschätzen Beobachtungsstudien den „wahren“ Effekt, wenn z.B. alle plausiblen Biasformen in die

Tab. 19.5: Faktoren, die die Qualität der Evidenz höher stufen

- Große Effekte bei direkter Evidenz, einem $RR > 2$ oder $RR < 0,5$ ohne plausible Confounder → Anheben um 1 Qualitätsstufe;
- sehr große Effekte mit einem $RR > 5$ oder $RR < 0,2$ ohne plausible Confounder → Anheben um 2 Qualitätsstufen
- Alle noch nicht berücksichtigten plausiblen Confounders haben den nachgewiesenen Effekt wahrscheinlich verkleinert → Anheben um 1 Qualitätsstufe
- Nachweis einer Dosis-Wirkungs-Beziehung → Anheben um 1 Qualitätsstufe

Richtung wirken, dass sie den scheinbaren Behandlungseffekt reduzieren. In der Regel sind wir insbesondere in Sorge, dass ein Bias einen tatsächlich vorhandenen Effekt fälschlicherweise vergrößert und suchen daher immer nach Faktoren, die zu einer Überschätzung des Effekts führen könnten und die in der Studie bisher nicht berücksichtigt wurden. Man kann sich Konstellationen vorstellen, bei denen ein (plausibler) Bias fälschlicherweise zu einer Abschwächung des Effekts, nicht jedoch zu einer Überschätzung des Effekts führen würde. Eine solche Konstellation entsteht, wenn z.B. nur kränkere Patienten eine Intervention (oder Exposition) erhalten, sie aber dennoch ein besseres Ergebnis hätten als die (weniger kranke) Kontrollgruppe. Dies würde die Schlussfolgerung nahelegen, dass der tatsächliche Effekt der Intervention (oder Exposition) größer ist als in der Studie beobachtet. So fanden Forscher zum Beispiel in einem systematischen Review aus Beobachtungsstudien an 38 Millionen Patienten in Privatkliniken höhere Mortalitätsraten als in vergleichbaren Not-for-profit-Kliniken [17]. Eine Biasform, die diesen Effekt erklären könnte, bezieht sich auf Unterschiede im Krankheitschweregrad der Patienten, der möglicherweise zwischen den beiden Kliniktypen existiert. Es ist plausibel, dass Patienten der Not-for-profit-Kliniken schwerer erkrankt sind als Patienten der privaten For-profit-Krankenhäuser. Für den Fall,

dass in den Daten ein verbleibendes Confounding existiert, würde sich der Bias gegen die Not-for-profit-Kliniken richten (d.h. in den Not-for-profit-Kliniken mit den kränkeren Patienten würde man die höheren Mortalitätsraten erwarten.) Da jede plausible Biasform in den Studien den beobachteten Behandlungseffekt reduzieren würde, kann man die Qualität dieser Beobachtungsstudien als moderat statt als niedrig erachten.

- ▲ Der Nachweis einer Dosis-Wirkungsbeziehung kann ebenfalls die Glaubwürdigkeit der Ergebnisse von Beobachtungsstudien verstärken und die zugeordnete Qualitätsstufe anheben. Beispielsweise zeigen Beobachtungsstudien bei Patienten mit zu hoher therapeutischer Marcumarisierung eine erhöhte Blutungsrate. Die Dosis-Wirkungsbeziehung zwischen höheren INR (International normalised ratio)-Werten und erhöhtem Blutungsrisiko legt nahe, dass der Effekt wirklich existiert.

19.1.7 Was tun, wenn sich die Qualität der Evidenz zwischen den verschiedenen Outcomes stark unterscheidet?

Handlungsempfehlungen gründen sich im Allgemeinen auf mehrere Outcomes und in dem oben beschriebenen Ansatz würde man für jeden patientenrelevanten Endpunkt die Qualität bestimmen. Hier kann ein Dilemma

entstehen: Wie soll man die Gesamtqualität der Evidenz festlegen, wenn sich die Qualität zwischen den einzelnen patientenrelevanten Endpunkten unterscheidet? Nehmen wir als Beispiel die selektive Darmdekontamination (SDD) bei Intensivpatienten. Mehrere hochwertige Meta-Analysen aus RCTs haben nachgewiesen, dass die SDD die Inzidenz von Infektionen reduziert und wahrscheinlich auch die Mortalität bei Beatmungspatienten mit SDD senkt. Die Qualität der Datenlage über die Auswirkung von SDD auf das Auftreten von Antibiotika-Resistenzen und seine klinische Relevanz hat dagegen wesentliche Mängel und man würde die Evidenz über diese gefürchtete seltene Nebenwirkung als niedrige Qualität einschätzen [18]. Soll man nun die Gesamtqualität der Evidenz über den Einsatz von SDD also als hoch, mittelwertig oder niedrig einschätzen?

In solchen Situationen schlägt das GRADE-System den Leitliniengruppen vor, abzuwägen, inwieweit die Nebenwirkungen der Therapie bei der Festlegung der optimalen Behandlungsstrategie einen entscheidenden Einfluss haben oder nicht. Ist die Qualität der Evidenz für einen Endpunkt, der eine Behandlungsstrategie entscheidend mitprägt, niedrig, muss die geringere Qualität dieses Endpunkts bei der Gesamteinschätzung der Qualität auch berücksichtigt werden. Ist der Endpunkt mit geringwertiger Evidenz zwar wichtig, aber nicht entscheidungstragend, dann soll sich die Gesamtbewertung auf die entscheidenden Endpunkte beziehen, für die Evidenz von hochwertiger Qualität vorliegt.

19.1.8 Wie sich mit dem GRADE-System die Evidenz abbilden lässt und Empfehlungen entwickeln lassen

Tabelle 19.6 enthält eine Checkliste, wie sich das GRADE-System bei der Entwicklung und der Graduierung von Empfehlungen einsetzen lässt.

Der gesamte Bewertungsprozess lässt sich in einer speziellen Software GRADE-Profilierer (GRADEpro, www.gradeworkinggroup.org) festhalten. Das Evidenzprofil beschreibt auf einer einzigen Seite die detaillierte Qualitätsbewertung, Wirksamkeit (bzw. Schaden) als relativen und absoluten (un-)erwünschten Effekt und die Bedeutung, die die Leitliniengruppe jedem einzelnen Endpunkt beigegeben hat. Damit werden die Ergebnisse der oft mehrmonatigen Arbeit für jeden entscheidenden Endpunkt nochmals knapp, aber umfassend visualisiert. Diese eine Seite umfassende Übersicht hat sich bei der Formulierung der Empfehlung als hilfreich erwiesen, wenn sich die Leitliniengruppen die Diskussionen und Einzelentscheidungen der vergangenen Monate nochmals in Erinnerung rufen müssen. Auf der Homepage der GRADE Working Group finden Sie ein entsprechendes Evidenzprofil für die Behandlung von Vorhofflimmern mit oralen Antikoagulantien.

Webmaterial

Zusätzliche Beispiele und Tabellen finden sich auf der Homepage der GRADE WORKING GROUP unter http://www.gradeworkinggroup.org/_DE/
Tabelle H1: Das GRADE-System für die Entwicklungen von Empfehlungen
Tabelle H2: Das GRADE-Evidenzprofil: Behandlung von Vorhofflimmern mit oralen Antikoagulantien (Beispiel).

19.2 Ausgewählte Aspekte aus der Leitlinienarbeit in Deutschland

19.2.1 Eine differenzierte Empfehlung braucht eine fokussierte Frage

Ausgangspunkt für jede Leitlinienempfehlung ist eine präzise Fragestellung, die auf einem Versorgungsproblem des Arztes beruht und für die Leitliniengruppe am Ende eine

Tab. 19.6: Checkliste für die Entwicklung und Einstufung von Empfehlungen

- Definition der Population, der Intervention und ihrer Behandlungsalternative(n) sowie der relevanten Endpunkte
- Festlegung der relativen Bedeutung der Endpunkte für die Entscheidungsfindung (entscheidungsstragend; klinisch wichtig, aber nicht entscheidungstragend, klinisch ohne wesentliche Bedeutung)
- Erstellung von systematischen Evidenzberichten (nach Möglichkeit auf der Grundlage von systematischen Reviews oder einer systematischen Bewertung der vorhandenen Evidenz)
- RCTs starten von einer hohen Qualitätsstufe, gut gemachte Beobachtungsstudien von einer niedrigen Qualitätsstufe, und es schließen sich folgende Schritte an:
 - Überprüfung auf schwerwiegende methodische Schwächen (fehlende Verblindung, fehlende verdeckte Randomisierung, hoher Verlust bei der Nachbeobachtung, vorzeitiger Studienabbruch wegen angeblich hohen Nutznachweises)
 - Indirekte Vergleiche bei Populationen, Interventionen oder Endpunkten (z.B. Surrogatparameter)
 - Inkonsistente Ergebnisse (Heterogenität)
 - Ungenaue Effektschätzungen (breite Konfidenzintervalle)

Je nach Ausprägungsgrad führen diese Mängel zu einer Abstufung der RCTs von hochwertiger zu mittlerer, niedriger oder sehr niedriger Qualität und bei Beobachtungsstudien zu einer Abstufung zu sehr niedriger Qualität.
- Liegen keine RCTs, sondern nur gut gemachte Beobachtungsstudien vor, startet die Bewertung bei niedriger Qualität. Falls ...
 - ein großer oder sehr großer Effekt vorliegt,
 - alle plausiblen Confounder den Effekt der Maßnahme nur verkleinern würden,
 - eine Dosis-Wirkungsbeziehung vorliegt,

kann man die Qualitätsstufe, je nach Ausprägung der Stärken, auf mittel- oder hochwertig erhöhen. Nur Beobachtungsstudien ohne Einschränkungen in der methodischen Qualität lassen sich höher stufen. Studien von sehr niedriger Qualität und Beobachtungsstudien mit methodischen Mängeln können nicht aufgewertet werden.
- Bestimmung der Effektgrößen für erwünschte und mögliche unerwünschte Wirkungen, der Belastungen und Kosten für die jeweilige Population
- Gesamtabwägung, inwieweit die erwünschten Wirkungen mögliche unerwünschte Wirkungen, Belastungen für den Patienten und Kosten für die relevante Population aufwiegen und wie eindeutig und schätzgenau diese Abwägung im Kontext der Werte und Vorstellungen der Patienten ist.

Handlungsempfehlung aussprechen möchte. Dieser erste Schritt hat Auswirkungen auf alle weiteren Phasen im Erstellungsprozess. Anhand der Frage wird die Literatursuche durchgeführt und die Kriterien für den Ein- und Ausschluss der Studien festgelegt. Je klarer durchdacht und fokussiert die Frage, desto größer ist die Wahrscheinlichkeit, eine nutzbare Antwort zu erhalten. Je weniger durchdacht und fokussiert die Frage, desto größer ist auch das Risiko, während der Studienauswertung erkennen zu müssen, dass die gesuchte Literatur die intendierte (aber nicht formulierte) Frage gar nicht erfasst. Die

Frage ist gut formuliert, wenn sie die Patientengruppe und ihr Gesundheitsproblem präzise definiert, die diagnostische oder therapeutische Option mit zugehöriger Alternative beschreibt und die Endpunkte benennt, die für den Patienten relevant sind.

Die Formulierung einer „suchtauglichen“ Frage aus einer klinischen Frage beinhaltet dabei bereits einen Übersetzungsprozess, bei dem die möglichst enge Zusammenarbeit zwischen Methodikern und Klinikern angestrebt werden sollte. Zum Beispiel wird eine Leitliniengruppe bei der unscharf formulierten und breit gefassten Frage: „Welche

Patienten sind für Maßnahme xy am besten geeignet?“ einen Berg an schwer sortierbaren Literaturzitataten, aber keine überzeugende Antwort erhalten. Eine suchtaugliche und beantwortbare Frage könnte lauten: „Um wie viel senkt die Maßnahme xy verglichen mit der Alternativbehandlung xyz die Mortalität und Morbidität bei Patienten mit Krankheit ab im Stadium cd?“ Darauf ließen sich weitere Fragen aufbauen, wie z.B.: „Gibt es einen Unterschied in den Endpunkten bei Patienten mit oder ohne Risikofaktoren, mit geringerem oder höherem Lebensalter?“ Auch die GRADE-Methodik verwendet fokussierte, strukturierte Schlüsselfragen als Ausgangspunkt für die Leitlinienentwicklung. Auf dieser Struktur ist auch das Evidenzprofil aufgebaut, das mit dem zugehörigen Programm (GRADE-Profilier) erstellt werden kann.

19.2.2 Der Begriff „klinische Relevanz“ lässt sich operationalisieren

Bereits mit der Wahl der Endpunkte in der Fragestellung legen Leitliniengruppen fest, welche Veränderungen im Gesundheitszustand der Patienten sie als maßgeblich betrachten. Dabei können sich die gewählten Endpunkte von den in Studien vorgefundenen Endpunkten unterscheiden, wenn Forscher in ihren Studien nur leicht messbare Parameter aus dem Krankheitsprozess gewählt haben [19], die aber aus klinischer Sicht nur begrenzte Rückschlüsse darüber erlauben, inwieweit die Intervention das Wohlergehen der Patienten auch wirklich verbessert.

Um die klinische Relevanz eines Endpunktes und seine Bedeutung für den Patienten besser einzuschätzen, schlägt die GRADE-Methodik vor, die Studien-Endpunkte eines Krankheitsbilds auf einer hierarchischen Skala mit den Ankerpunkten 1 (= klinisch unbedeutsam) und 9 (= von höchster klinischer Bedeutsamkeit) anzuordnen. Abbildung 19.2

zeigt am Beispiel der enteralen Ernährung geriatrischer Patienten zur verbesserten Heilung von Dekubitalulzera, wie eine Hierarchie der (Studien-)Endpunkte aussehen könnte. Je größer die Bedeutsamkeit für den Patienten, desto wichtiger werden die Endpunkte für die Formulierung der Empfehlung. Die hierarchische Visualisierung aller möglichen Endpunkte erleichtert einer Leitliniengruppe die Konsensfindung darüber, welche positiven Effekte einer Intervention nachgewiesen sein sollten, bevor man eine handlungsleitende Empfehlung ausspricht.

Die Frage nach dem kleinsten klinisch relevanten Unterschied zwischen einer Intervention und der Kontrollgruppe ist ein weiterer häufiger Diskussionspunkt in Leitliniengruppen. Sie ist schwierig zu beantworten, denn dieser kleinste relevante Unterschied definiert sich stark über den Kontext der Fragestellung, den Schweregrad der Erkrankung, die Wirksamkeit alternativer Behandlungsformen und das Ausgangsrisiko des Patienten, ohne Behandlung einen unerwünschten Endpunkt zu erleiden. Ein Lösungsansatz wäre, die beobachteten Effekte im Verhältnis zu anderen Interventionen oder Krankheitsbildern in „große eindeutige Effekte“ und „kleine marginale Effekte“ einzuteilen und bei der Entscheidung die relativen wie die absoluten Effekte und die Präzision der Schätzung (das dazugehörige Konfidenzintervall) heranzuziehen. Nach der GRADE-Methodik verdienen große Effekte mit engem Konfidenzintervall eher eine starke Empfehlung als kleine Effekte mit einem weiten Konfidenzintervall. Parallel dazu sollten die unerwünschten Effekte und die Belastungen für den Patienten (und gegebenenfalls das Kosten-Nutzen-Verhältnis der Intervention) zusammengetragen und den erwünschten Effekten gegenübergestellt werden. Erst die Gesamtschau der Vor- und Nachteile ermöglicht es Leitliniengruppen, eine kontextbezogene Wertung darüber abzugeben, ob die Größe eines beobachteten

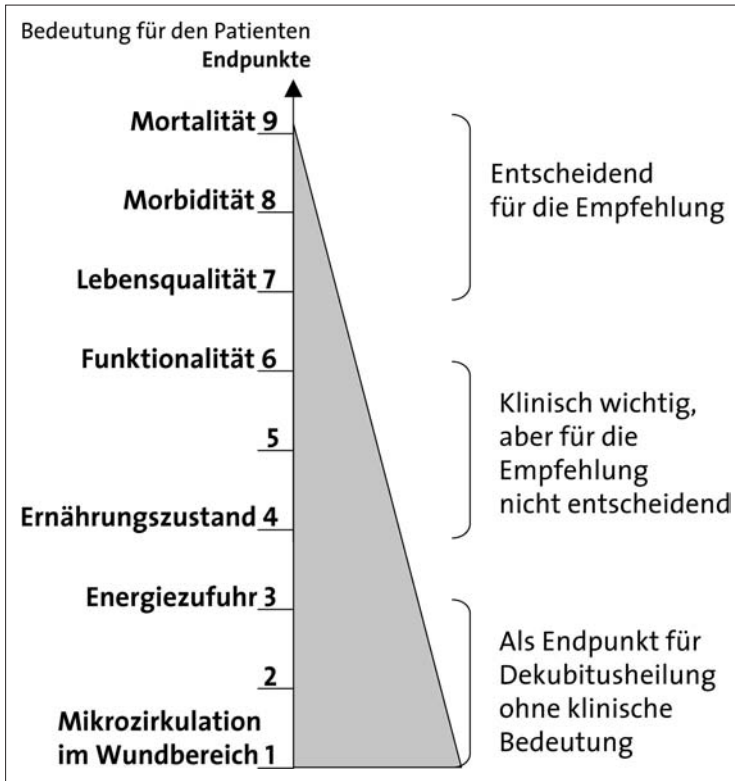


Abb. 19.2: Hierarchie der Endpunkte nach ihrer Bedeutung für den Patienten am Beispiel der enteralen Ernährung geriatrischer Patienten zur Dekubitusheilung

Effekts noch klinisch relevant ist oder ob man nach Abwägung der Vor- und Nachteile einen kleinen Effekt als klinisch irrelevant erklärt, um aus dieser Zusammenfassung dann eine Empfehlung abzuleiten.

19.2.3 Nicht jede Aussage resultiert in einer Empfehlung

Evidenz kann nur den aktuellen Wissensstand abbilden, d.h. Evidenz führt nicht automatisch zu einer Handlungsempfehlung [20]. Vielmehr muss sich die Leitliniengruppe bei der Auswahl der Schlüsselfragen überlegen, für welche Situationen Ärzte und Patienten eine Empfehlung über den aktuellen Standard der Versorgung benötigen und eine Standardsetzung als hilfreiche Orientierung empfinden. Empfehlungen sind handlungsleitend und wollen eine Aktion in Gang set-

zen (oder diese verhindern). Aufgearbeitetes „Faktenwissen“ (z.B. Prävalenzangaben, Klassifikationen oder Schweregradeinteilungen) ohne Handlungskonsequenz kann in einem „Statement“ mit Evidenzstufe resultieren, nicht aber in einer Empfehlung. Es bedarf daher auch keiner Empfehlungsgraduierung.

19.2.4 Die Empfehlung leitet sich aus der Evidenz ab

Die Dissoziation zwischen Evidenz und Empfehlung war eine typische Schwäche der früheren Expertenempfehlung (vgl. Abb. 19.3) [21] und man glaubte, mit dem evidenzbasierten Ansatz diesen Mangel überwunden zu haben. Ein kritischer Blick in aktuelle Leitlinien zeigt, dass diese Hoffnung verfrüht war. Vielmehr entwickelt sich eine

neue Variante der Expertenleitlinie, bei der die Evidenz nach anspruchsvollen methodischen Regeln aufgearbeitet wird, die Leitliniengruppe diese Evidenz in der Formulierung der Empfehlung aber unzureichend berücksichtigt. So beantwortete z.B. die Deutsche Gesellschaft für Ernährungsmedizin die Frage „2.9. Lassen sich bei multimorbiden Patienten durch enterale Ernährung Dekubitalulzera vermeiden bzw. verbessern?“ wie folgt: „Aufgrund eindeutig positiver klinischer Erfahrungen wird die enterale Ernährung zur Verbesserung von Dekubitalulzera empfohlen“ [22]. Die sorgfältig aufgearbeiteten randomisierten und nicht randomisierten Studien lieferten jedoch keinen Beleg für die Empfehlung. Im Gegenteil: konsistente ernstzunehmende Hinweise für eine erhöhte Rate an Aspirationspneumonien und eine er-

höhte Mortalität in der Behandlungsgruppe stellen die Harmlosigkeit der enteralen Ernährung in Frage. Auch die häufig auftretende Notwendigkeit der Fixierung von Patienten, die mit Nasensonde ernährt werden müssen, und die damit einhergehende Beeinträchtigung der Lebensqualität finden keinen Eingang in die Empfehlung.

Eine ähnliche Situation findet sich bei den Empfehlungen zur Selbstuntersuchung der Frau in der S3-Leitlinie Brustkrebs-Früherkennung [23]: Im Hintergrundtext der Leitlinie (4.4. Selbstuntersuchung, S. 51) wird auf Studienbasis die fehlende Wirksamkeit der Brust-Selbstuntersuchung dokumentiert (keine Auswirkung auf die Mortalität durch Mammakarzinom oder das Auffinden von Tumoren in einem Frühstadium) und die potenziellen Nachteile der Selbstuntersu-

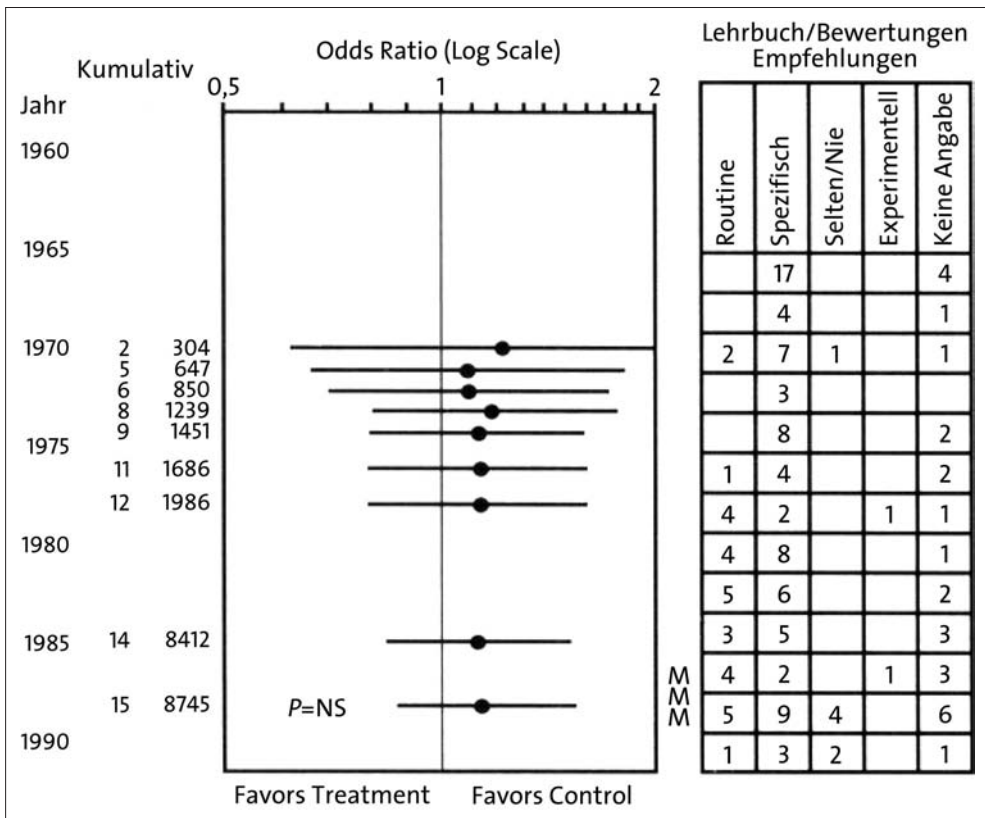


Abb. 19.3: Effekt der prophylaktischen Lidocaingabe auf die Mortalität bei Patienten mit akutem Herzinfarkt – Diskrepanz zwischen wissenschaftlicher Datenlage und Handlungsempfehlungen [21]

chung explizit beschrieben (psychische Belastung durch zyklusbedingte Gewebeveränderungen, falsch positive Befunde, Überbewertung als Maßnahme der Früherkennung). Die resultierende Empfehlung formuliert jedoch eine uneingeschränkt positive Aussage zugunsten der Selbstuntersuchung („... muss die Selbstuntersuchung Bestandteil eines Früherkennungsprogramms sein und bleiben ...“.)

In beiden Fällen zeigt sich eine Dissoziation zwischen der zitierten Studienevidenz und der abgegebenen Empfehlung. Werden die Gründe nicht dargelegt, ist es den Anwendern der Leitlinie nicht möglich, den Diskussionsprozess, der zu dieser Abweichung geführt hat, nachzuvollziehen, und es kann ein Zweifel an der methodischen Qualität der Leitlinie entstehen. Anstatt zu versuchen, die Diskrepanz zwischen Datenlage und Meinung transparent abzubilden und aufzulösen, wird die Evidenz bei der Formulierung der Empfehlung ignoriert und eine den Nutzern – ÄrztInnen und Frauen – nicht nachvollziehbare Empfehlung ausgesprochen. Ein Vorgehen entsprechend der GRADE-Methodik kann zur geforderten Transparenz beitragen und wird in Leitliniengruppen möglicherweise zu einer stärkeren Berücksichtigung der Studiendaten führen.

19.2.5 Zusammenfassung

Mit der GRADE-Methodik wurde ein System entwickelt, das vermehrt Transparenz und Reproduzierbarkeit in die Erstellung von Leitlinien bringt. Dabei kommt der Identifikation und Offenlegung der Werte und Vorstellungen, die den Empfehlungen zugrunde liegen, ein besonderer Stellenwert zu. Die GRADE-Methodik bietet auch Leitliniengruppen in Deutschland Lösungsansätze für kritische Fragen in der Leitlinienentwicklung, insbesondere der Formulierung und Graduierung von Empfehlungen.

Literatur

- [1] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI). Fassung 2005/2006. Z Arztl Fortbild Qualitatssich. (2005), 99, 468–519
- [2] Schunemann HJ et al., An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. Am J Respir Crit Care Med. (2006), 174, 605–614
- [3] Guyatt G et al., Grading strength of recommendations and quality of evidence in clinical guidelines: report from an american college of chest physicians task force. Chest. (2006), 129, 174–181
- [4] Schunemann H et al. (2007) Grading Recommendations. In: Guyatt G, Rennie D, Users' Guide to the Medical Literature: A Manual for Evidence-Based Clinical Practice
- [5] Sackett DL, Rules of evidence and clinical recommendations on the use of anti-thrombotic agents. Chest. (1986), 89, 2S–3S
- [6] WHO rapid advice guidelines on pharmacological management of humans infected with avian influenza (H5N1) virus. http://www.who.int/csr/disease/avian_influenza/guidelines/pharmamanagement/en/index.html (01.12.06)
- [7] Uhlig K et al., Grading evidence and recommendations for clinical practice guidelines in nephrology. A position statement from Kidney Disease: Improving Global Outcomes (KDIGO). Kidney Int. (2006), 70, 2058–2065
- [8] Rose BD, Graded Recommendations now in UpToDate (14.1 News). http://www.uptodate.com/img/14_1_newsletter.pdf (01.12.06)
- [9] Atkins D et al., Grading quality of evidence and strength of recommendations. BMJ. (2004), 328, 1490–1497
- [10] Schunemann HJ et al., Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. CMAJ. (2003), 169, 677–680
- [11] van Walraven C et al., Oral anticoagulants vs aspirin in nonvalvular atrial fibrillation

- on: an individual patient meta-analysis. *JAMA*. (2002), 288, 2441–2448
- [12] Devereaux PJ et al., Differences between perspectives of physicians and patients on anticoagulation in patients with atrial fibrillation: observational study. *BMJ*. (2001), 323, 1218–1222
- [13] Guyatt G et al., Addressing resource allocation issues in recommendations from clinical practice guideline panels: suggestions from an American College of Chest Physicians task force. *Chest*. (2006), 129, 182–187
- [14] Bhandari M et al., Association between industry funding and statistically significant pro-industry findings in medical and surgical randomized trials. *CMAJ*. (2004), 170, 477–480
- [15] Alonso-Coello P et al., Meta-analysis of flavonoids for the treatment of haemorrhoids. *Br J Surg*. (2006), 93, 909–920
- [16] Cannegieter SC, Rosendaal FR, Briet E, Thromboembolic and bleeding complications in patients with mechanical heart valve prostheses. *Circulation*. (1994), 89, 635–641
- [17] Devereaux PJ et al., A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals. *CMAJ*. (2002), 166, 1399–1406
- [18] Garnacho-Montero J et al., Impact of adequate empirical antibiotic therapy on the outcome of patients admitted to the intensive care unit with sepsis. *Crit Care Med*. (2003), 31, 2742–2751
- [19] Bucher HC et al. (2007) *Surrrogate Outcomes*. In: Guyatt G, Rennie D, *Users' Guide to the Medical Literature: A Manual for Evidence-Based Clinical Practice*
- [20] Haynes RB, Devereaux PJ, Guyatt GH, Physicians' and patients' choices in evidence based practice. *BMJ*. (2002), 324, 1350
- [21] Antman EM et al., A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. *Treatments for myocardial infarction*. *JAMA*. (1992), 268, 240–248
- [22] Leitlinie Enterale Ernährung der DGEM und DGG: Enterale Ernährung (Trink- und Sondennahrung) in der Geriatrie und geriatrischen neurologischen Rehabilitation. *Aktuel Ernährungsmed*. (2004), 29, 198–225
- [23] Schulz KD, Albert US, *Konzertierte Aktion Brustkrebs-Früherkennung in Deutschland (2004) Stufe-3-Leitlinie*. Brustkrebs-Früherkennung in Deutschland. Zuckschwerdt, München

20 Wo erhalte ich Antwort auf meine Fragen?

Monika Lelgemann, Norbert Donner-Banzhoff

20.1 Einleitung

Für Ärztinnen und Ärzte, die heute Antworten auf ihre Fragen suchen, ist das „Zuviel“ eher ein Problem als das „Zuwenig“. Unter vielen anderen die einschlägige Quelle zu finden, die interessierende Information von irrelevantem Material zu trennen, die Hintergründe und Interessen medizinischer Medien zu erkennen – dies umschreibt die Grundaufgaben ärztlichen Informationsmanagements heute.

Hier kommen wir zum Kern der EbM: die eigenen Fragen ernst nehmen und nach Antworten suchen. Dies setzt voraus, verschiedene Informationsquellen zu kennen, einzuschätzen und kompetent zu recherchieren. Dabei läuft im Hintergrund immer eine Sanduhr mit: jeder Schritt einer evidenzbasierten Recherche steht unter dem Diktat der Zeitökonomie. Zu den Grundsätzen und Verdiensten der Evidenzbasierten Medizin gehört es, ein Ordnungsprinzip für den „Dschungel“ der Information anzubieten, das sich in erster Line an methodischen Aspekten orientiert.

Allerdings ist bekannt, dass gegenüber der Umsetzung dieser Prinzipien der Evidenzbasierten Medizin in der Praxis erhebliche Barrieren bestehen. Hierzu gehören unter anderem Zeitmangel, eingeschränkter

Zugang (Technik, Kosten) zu relevanten Quellen, unübersichtliche, nicht einheitliche Datenbanken, deren Benutzung jeweils eine spezifische Einarbeitung erforderlich macht sowie eine nutzerunfreundliche Darstellung der gefundenen Quellen [1; 2]. Ein weiterer Aspekt besteht darin, dass viele Ärzte angeben, die gefundene Literatur hinsichtlich ihrer Validität nicht ausreichend beurteilen zu können.

20.2 Eine schwierige Abwägung

Aus den oben genannten Gründen sind Ärztinnen und Ärzte, die Antworten auf ihre Fragen suchen, überwiegend auf Quellen aufbereiteter Evidenz angewiesen [3]. Darüber hinaus müssen sie bei jedem Schritt ihrer Informationssuche überlegen, ob der zu erwartende Wissenszuwachs, das „Mehr“ an Information, den Aufwand rechtfertigt bzw. ob das zusätzliche Wissen entscheidungsrelevant werden wird.

Wenn wir Informationen suchen und bewerten, müssen wir sich widersprechenden Gesichtspunkten gerecht werden. Das entsprechende Abwägen können wir uns mit der didaktischen Formel von Slawson und Shaughnessy [4] verdeutlichen:

Abb. 20.1: Nutzen einer Information

$$\text{Nutzen einer Information} = \frac{\text{Relevanz} \times \text{Validität}}{\text{Aufwand}}$$

20.2.1 Relevanz

Hierzu ist der schöne Begriff des POEM (Patient orientated evidence that matters) geprägt worden. Darunter verstehen wir Informationen, die

- ▲ eine relevante klinische Frage ansprechen,
- ▲ patientenorientierte Outcomes messen (Morbidität, Mortalität, Beschwerden, Lebensqualität usw.) und
- ▲ deren Umsetzung zu einer Änderung der täglichen Praxis führt.

Im Gegensatz dazu steht die so genannte Disease orientated evidence (DOE), die sich auf Pathologie, Pathophysiologie u.Ä. bezieht, meist aber nicht zur Beantwortung der konkreten klinischen Frage führt. Die Unterscheidung zwischen DOEs und POEMs (s. a. Kap. 4.3) ist nicht nur bei konkreten Fragen („hunting“), sondern auch beim Durchstreifen („foraging“) der Literatur (z.B. regelmäßig gelesener Zeitschriften) nützlich. Sie erlaubt es, DOEs schnell zu verwerfen; dies kann klassischer Weise dann gefahrlos geschehen, wenn eine Intervention lediglich durch Surrogatkriterien evaluiert wird oder vom Leser sowieso nicht umsetzbar ist. So wurde in über 8000 Originalarbeiten lediglich in 2,6% ein POEM gefunden [5]. Dies ist natürlich ein ernüchterndes Resultat, andererseits stellt die geringe Menge **relevanter** Literatur wiederum eine Erleichterung für in Informationsfluten fast untergehende Ärzte dar.

20.2.2 Validität

Hierbei geht es um die Beurteilung der Vertrauenswürdigkeit einer Information. Das klassische Vorgehen der EbM versteht hierunter den Prozess der kritischen Bewertung z.B. einer Therapie-Studie an Hand des Studiendesigns. Diese kritische Bewertung einer Studie, oft mithilfe von Checklisten, stellt ei-

nes der Kernelemente der Umsetzung der Prinzipien der Evidenzbasierten Medizin dar.

Wenn für die tägliche Praxis jedoch die Zeit, vielleicht auch die methodischen Kenntnisse fehlen, ist ein vereinfachtes Verfahren erforderlich. So muss vielfach nach der Vertrauenswürdigkeit einer Quelle entschieden werden, ob eine Angabe akzeptiert wird oder nicht. Hier spielen persönliche Gesichtspunkte eine Rolle (Wertschätzung eines bekannten Kollegen), aber auch die Zugehörigkeit zu einer Fachgruppe (Leitlinien der eigenen Fachgesellschaft) oder zu einem Versorgungssektor (z.B. Misstrauen niedergelassener Ärzte gegen „die Klinik“) [6].

Dennoch sollte es zu einem professionellen Informationsmanagement gehören, sich bestimmte methodische Mindeststandards in der Beurteilung der Verlässlichkeit (Validität) von Quellen zu setzen. Neben dem klassischen Instrumentarium der EbM sind „schnelle Heuristiken“ zu entwickeln, die eine Sortierung und Bewertung ermöglichen [7].

20.2.3 Aufwand

Von entscheidender Bedeutung ist der Nenner. Eine Information mag noch so relevant und valide sein; wenn der Aufwand, sie zu finden, als zu hoch eingeschätzt wird, dann wird sie ein ungehobener Schatz bleiben. An erster Stelle sind hier die Zugänglichkeit von Datenbanken (Kosten, technische Voraussetzungen), ihre Struktur und Nutzeroberfläche zu nennen. Ob Anbieter evidenzbasierter Informationen es schaffen werden, mit technischen Lösungen den Aufwand zu minimieren und damit die „Schätze“ der EbM leichter auffindbar zu machen, bleibt abzuwarten.

Wie bei der Verschreibung von Medikamenten empfiehlt es sich, dauerhaft eine begrenzte Zahl von Quellen intensiv zu nutzen, um deren Stärken und Schwächen genau kennen zu lernen und zu nutzen. Auch hier gilt: „Weniger ist mehr.“

20.2.4 Nutzen

Die drei Faktoren auf der rechten Seite der Gleichung in Abbildung 20.1 stehen miteinander in einem Zusammenhang. Das Streben nach hoher, d.h. transparenter Validität erhöht den Aufwand; bei der sorgfältigen Lektüre von Originalarbeiten sind hier oft mehrere Stunden zu veranschlagen. Wenn ich den Aufwand dadurch reduziere, dass ich aufbereitete Quellen nutze, liefere ich mich deren Bias aus und muss mich unter Umständen mit einer geringen Validität begnügen.

20.3 Systematische Aufbereitung: die Evidenzpyramide

Haynes [8] hat eine hierarchische Pyramide spezifischer evidenzbasierter Quellen definiert.

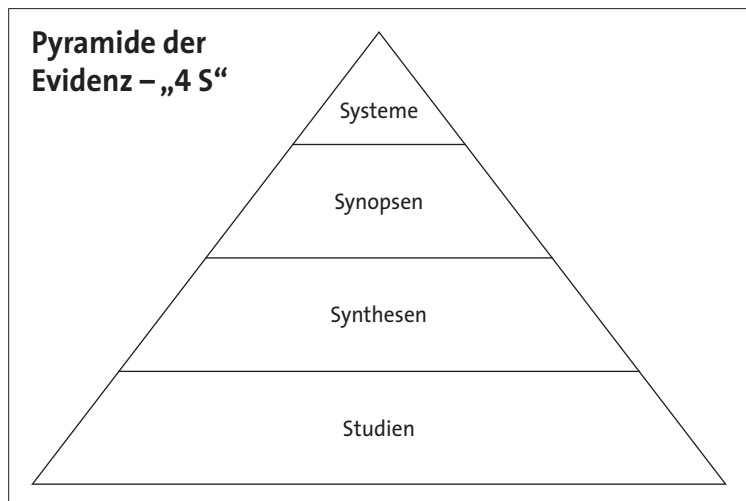
An der Spitze stehen „Systeme“, die im Idealfall in Routine-Dokumentationssysteme (Praxis, Krankenhaus) integriert sein sollten. Diese sollen die benötigte evidenzbasierte Information am „Punkt der Entscheidung“ liefern und jederzeit zur Verfügung stehen. Idealerweise sind sie mit automatischen Routinen ausgestattet, die ohne aktives Nachfragen auf-

grund von bestimmten Faktoren (z.B. Triggerdiagnosen) wichtige Hintergrundinformationen oder Handlungsoptionen geben.

Viele Leitliniengruppen, aber auch Betreiber von Datenbanken und anderen Informationssystemen bemühen sich derzeit um die Entwicklung und Integration solcher Systeme in die Software von Praxen und Krankenhäusern. Ihr Effekt im Hinblick auf Verhaltensänderungen von Ärzten ist empirisch belegt [9]. Auch elektronische Lehrbücher wie z.B. Clinical Evidence (s. unten) entwickeln sich in die Richtung umfassender, flexibler und gut aufbereiteter Systeme.

Kann auf der Stufe der Systeme die gewünschte Information nicht gefunden werden bzw. stehen diese nicht zur Verfügung, ist die Ebene der „Synopsen“ aufzusuchen. Hierunter versteht Haynes strukturierte Kurz-Zusammenfassungen von klinischen Studien und Synthesen, wie sie z.B. der ACP Journal Club und Evidence-based Medicine regelmäßig veröffentlichen. Diese sind in Papierform und als elektronische Datenbanken verfügbar. Neben strukturierten Zusammenfassungen der Primärarbeiten enthalten sie auch Kommentare klinischer bzw. methodischer Begutachter, die dem Anwender die Einordnung und Interpretation der Studienergebnisse erleichtern sollen.

Abb. 20.2: Pyramide der Evidenz



Auf Arzneimittel bezogen liefern in Deutschland das arznei-telegramm und der Arzneimittelbrief einen ähnlichen Service, ebenso wie der Evimed-Journalclub des Horton-Zentrums in Zürich. Dieses bietet nach Fachgebieten geordnet strukturierte Zusammenfassungen wesentlicher Primärstudien und systematischer Übersichtsarbeiten.

Als weitere Möglichkeit stehen die „Synthesen“ zur Verfügung. Klassisches Beispiel ist die Cochrane Library mit ihren nach festen Protokollen erstellten systematischen Übersichtsarbeiten (s. Kap. 43). Aber auch HTA (Health Technology Assessment)-Berichte und so genannte Evidenzberichte können zu diesen Quellen gezählt werden. Dabei ist zu beachten, dass sich HTA-Berichte in erster Linie an Entscheidungsträger innerhalb des Gesundheitssystems und nicht an praktizierende Ärzte richten.

Erst an vierter Stelle kommen die „Studien“, also der Bezugspunkt der „klassischen“

EbM, die z.B. über MEDLINE identifiziert werden können. Originalarbeiten sind für in der Versorgung tätige Ärzte nicht der bevorzugte Weg zur Beantwortung von Fragen, da hierbei Aufwand und Nutzen meist nicht in einem vertretbaren Verhältnis stehen [10].

20.4 Der „Weg des Wissens“

Die zuvor genannte Pyramide orientiert sich am Transparenz-Ideal der EbM. Ihre Etagen sind z.T. noch im Rohbau, der Zielgruppe sind die Möglichkeiten oft noch nicht bekannt. Trotzdem wird sehr viel Wissen aus klinischen Studien aufbereitet und an potenzielle Nutzer weitergeleitet, z.B. an verschreibende Ärzte. Auf diesem „Weg des Wissens“ finden wir allerdings zahlreiche Windungen und Verzerrungen.

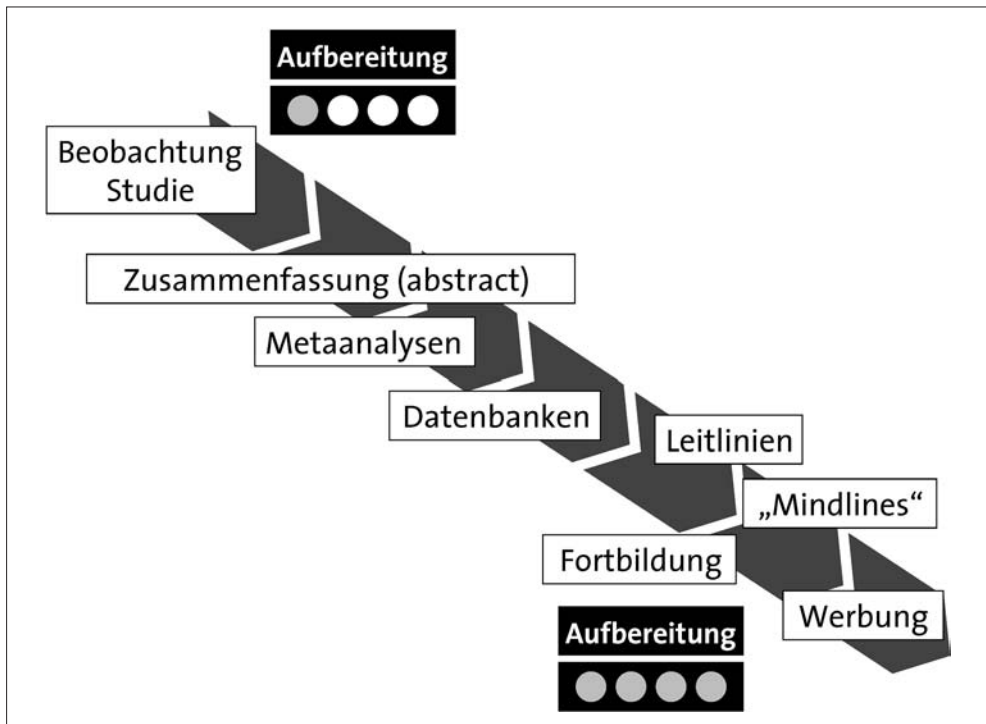


Abb. 20.3: Weg des Wissens

20.4.1 Studien

In der „rohesten“ Form liegt medizinisches Wissen als ärztliche Beobachtung vor; erfolgt diese nach wissenschaftlichen Regeln, sprechen wir von einer Studie. Doch schon die Darlegung der Ergebnisse in einer wissenschaftlichen Veröffentlichung bedeutet eine Informationsreduktion, verlangt Annahmen und die Entscheidung zwischen verschiedenen methodischen Strategien.

Dies gilt in noch stärkerem Maße für die Zusammenfassung (Abstract) einer wissenschaftlichen Arbeit. Diese enthält oft schon eine Tendenz, die die Daten in einem bestimmten Licht erscheinen lässt [11]. Trotzdem ist die Zusammenfassung einer Studie für eine schnelle Orientierung unverzichtbar. Einen großen Vorteil stellt auch der schnelle und kostenlose Zugang zumindest zu den Abstracts der in MEDLINE (s. auch Anhang) gelisteten Zeitschriften dar. Trotzdem sind Abstracts bzw. Originalarbeiten meist keine realistische Option, um (massen-) wirksam Wissen zu transportieren.

20.4.2 Systematische Übersichtsarbeiten

Systematische Übersichtsarbeiten orientieren sich eng an den Originalarbeiten und bieten grundsätzlich die Möglichkeit, subjektive Einflüsse zu minimieren. In einer der Datenbanken der Cochrane-Library, der Cochrane Database of Systematic Reviews, sind ca. 1500 systematischer Übersichtsarbeiten enthalten, die nach einem einheitlichen Protokoll erarbeitet worden sind (s. Kap. 43).

Großer Vorteil dieser Zusammenstellung ist unter anderem die einheitliche Erarbeitungsweise und Darstellung der Einzelpublikationen. Für Ärzte in Deutschland ist allerdings der Zugang zu dieser Datenbank kostenpflichtig, sodass die Informationen häufig nicht zu Verfügung stehen. Die Abs-

tracts der systematischen Übersichtsarbeiten sind jedoch kostenlos zugänglich.

Zu berücksichtigen ist ebenfalls, dass sich bislang viele der auftretenden Fragen nicht durch eine Recherche in der Cochrane Library beantworten lassen.

20.4.3 Clinical Evidence

Hierbei handelt es sich um den nächsten Schritt hin zu aufbereiteter und damit zwangsläufig „gewerteter Evidenz“. Clinical Evidence liegt als Web-gestützte Datenbank, Druckausgabe und als „Kompendium Evidenzbasierte Medizin“ in deutscher Übersetzung vor. Zu einem breiten Spektrum von Problemen bzw. Erkrankungen wird das Wissen aus randomisierten kontrollierten Studien und systematischen Übersichten referiert, und zwar überwiegend zu Fragestellungen aus Therapie und Prävention. Über den Stand der Evidenz aus (meistens) randomisierten kontrollierten Studien wird nicht hinausgegangen.

Ein gewichtiger Vorteil ist die deutschsprachige Ausgabe; einzelne Kapitel sind von deutschsprachigen Experten insbesondere im Hinblick auf Besonderheiten des nationalen Gesundheitssystems kommentiert. Diese Quelle muss als Buch bzw. Abonnement erworben werden, ein Zugang zur Web-Datenbank ist dabei enthalten.

20.4.4 Leitlinien

Leitlinien gehen über die Aufbereitung von Evidenz hinaus und stellen ein vollständiges Handlungsparadigma dar (s. a. Kap. 15). Hier ist inzwischen die Orientierung an Studien ein „Muss“, das allerdings in Deutschland noch längst nicht alle Leitlinien erfüllen. In den in Leitlinien enthaltenen Empfehlungen spielen immer auch andere Aspekte eine Rolle, wie z.B. ethische Überlegungen, Traditio-

nen oder klinische Erfahrungen. Leitlinien werten in höherem Maße als systematische Übersichtsarbeiten oder gar Primärstudien. Im Fall von methodisch hochwertigen Leitlinien sollte der Entwicklungsprozess von einer repräsentativen Gruppe beteiligter Ärzte und Vertretern anderer Berufsgruppen in einem formalisierten Verfahren erfolgt sein, sodass eine zu starke Einflussnahme Einzelner oder einzelner Interessengruppen minimiert werden kann.

Die meisten Leitlinien sind im Internet frei verfügbar. Gute Leitlinien enthalten zusätzliche Materialien, die z.B. die Vermittlung von Wissen an Patienten erleichtern können (Umsetzungshilfen).

20.4.5 Fortbildung – Mindlines – Faustregeln

Obwohl Leitlinien eine hoch aufbereitete Form der Evidenz bieten, dürfte nur eine Minderheit der Ärzteschaft sie intensiv studieren. Die meisten verlassen sich dann doch auf noch leichter zugängliche Formen, z.B. einen gut zu lesenden Fortbildungsartikel. Dass sich hier weitere Möglichkeiten des Bias ergeben, ist offensichtlich. So sind kostenlos versandte „Streuzeitschriften“ von inserierenden Herstellern abhängig und können sich eine kritische Berichterstattung überhaupt nicht leisten. Örtliche „Meinungsführer“ werden von der Industrie systematisch zur Beeinflussung der Kollegen-schaft eingesetzt.

Bei einer systematischen Befragung britischer Allgemeinärzte hat sich ergeben, dass „mindlines“ eine große Rolle spielen [12]. Dies sind lokal vorhandene „Weisheiten“, die sich durch eine spezielle Arbeitsteilung ergeben. Ein Kollege in der Gruppenpraxis oder im Qualitätszirkel hat sich mit einem Thema intensiv auseinandergesetzt; seine Schlussfolgerung wird – entsprechendes Vertrauen vorausgesetzt – von den Kollegen,

evtl. nach Diskussion, übernommen. Die Berge von Leitlinien, die die befragten Kollegen bzw. Praxen unaufgefordert erhalten, erzeugen dagegen ein Gefühl von Machtlosigkeit und Frustration; sie werden eigentlich nur von wenigen Kollegen bei der Konzeption der genannten Mindlines näher beachtet.

Etwas anders sind die „Faustregeln“ geteilt, die eine schwedische Allgemeinärztin mit Hilfe mehrerer Fokusgruppen von Allgemeinärzten identifiziert hat [13]. Hierbei handelt es sich um klinische Erfahrungsregeln, die sich vom in Lehrbüchern dargelegten expliziten Wissen oft deutlich unterscheiden.

Diese Formen von Wissen sind eine Realität, die vor dem Hintergrund des Transparenzideals der EbM eher als „schmuddelig“ empfunden wird. Sie sind aber auch positiv als eine Hilfe zu verstehen: in einem Meer von Information bieten sie Orientierung auch für solche Themen, mit denen man sich aus Zeit-, Kosten- oder Kompetenzgründen nicht näher befassen kann.

20.5 Bias überall

Die EbM hat zwar ein hochentwickeltes Instrumentarium zum Aufdecken von Bias entwickelt, sie hat aber selten thematisiert, welche Interessen denn nun zu einer Verzerrung von Information führen. Die Wahrscheinlichkeit, am Ende des oben gezeichneten Weges beim „Endverbraucher“ anzukommen, ist nicht für alle Informationen gleich. Ob Studienresultate überhaupt veröffentlicht werden, ob sie ihren Weg in „Streuzeitschriften“ oder die Agenda von Fortbildungsveranstaltungen finden, wird wesentlich von Herstellern und anderen mächtigen Gruppierungen bestimmt.

20.6 Zusammenfassung

Eine Orientierung im Dschungel der Information bieten die folgenden Grundregeln:

- ▲ Überlegen Sie sich zu Beginn einer Recherche, welche Art von Information Sie benötigen. Nur wenn der Kurs klar ist, werden Sie sich nicht ablenken oder verwirren lassen von all den Nebeninformationen, die am Rande des Suchweges liegen.
- ▲ Steuern Sie ganz gezielt einschlägige Datenquellen an (Studie, Leitlinie, Synopse usw.). Legen Sie sich im Laufe der Zeit ein begrenztes Repertoire von Quellen an, die Sie gut einschätzen können. In den Tabellen am Ende dieses Kapitels ist eine Auswahl aufgeführt, die auf die Bedürfnisse der allgemeinärztlichen Praxis zugeschnitten ist.

Suchen Sie gezielt in den jeweils passenden Quellen: Eine Suche nach Leitlinien wird in MEDLINE wenig Erfolg bringen, da Leitlinien überwiegend als elektronische Dokumente zur Verfügung stehen und nicht in Zeitschriften veröffentlicht werden. Leitlinien lassen sich aber sehr wohl über eine der großen Internetsuchmaschinen (z.B. google) finden, wohingegen eine Suche nach Primärstudien über Internetsuchmaschinen wenig erfolgversprechend ist.

- ▲ Entscheiden Sie sich bewusst für eine sensitive oder spezifische Suche. Erstere will möglichst wenig Information verpassen, ist aber für Alltagsfragen meist zu aufwendig. Bei letzterer kann man sich wiederum durch zufällig gefundene Information (z.B. eine einzelne Studie) täuschen lassen. Hierin liegt eines der Hauptargumente für die Verwendung zusammengefasster Evidenz (z.B. systematische Übersichtsarbeiten), da so eine Verzerrung durch die Verwendung so genannter „opportunistischer Evidenz“ weniger wahrscheinlich wird.

- ▲ Machen Sie sich Ihren eigenen Bias und Ihr eigenes Erkenntnisinteresse klar. Haben Sie den Mut, gerade solche Evidenz zu suchen, die in eine andere Richtung weist.

20.7 Ausblick

Es fällt schwer, die künftige Entwicklung möglicher Quellen zur evidenzbasierten Beantwortung von Fragen abzuschätzen. Nach der ersten EbM-Dekade sehen wir die Notwendigkeiten und Möglichkeiten realistischer als in der ersten Euphorie. Manches methodisches Wissen wird Allgemeingut werden, wie z.B. die Unterscheidung von Absoluter und Relativer Risiko-Reduktion, die Wertschätzung klinisch relevanter Zielkriterien (Endpunkte) oder die Interpretation eines Konfidenzintervalls. Ansonsten bedürfen klinische Studien jedoch der Aufbereitung, um die Mehrzahl der Ärzte zu erreichen, die methodisch nicht speziell trainiert sind.

Die von Haynes avisierten „Systeme“ bedürfen dabei nicht nur einer intelligenten und benutzerfreundlichen Programmierung, sondern auch eines immensen inhaltlichen und methodischen Inputs. Die Cochrane Library stellt ein geglücktes Beispiel für eine nicht-kommerzielle Initiative zur Lösung dieses Problems dar. Hier wurden weltweit Hunderte von Wissenschaftlern mobilisiert, um ein riesiges Material an Studienergebnissen systematisch aufzuarbeiten. Ob es gelingen wird, Informations-Systeme entsprechend weiterzuentwickeln und von Hersteller-Interessen weitgehend freizuhalten, wird wesentlich vom Engagement der Ärzteschaft selbst abhängen. Dies betrifft einerseits die kritische Sichtung und Darstellung der Studienevidenz. Auf der anderen Seite müssen aber auch die „Konsumenten“ bereit sein, den Zugang zu aufbereiteten und verlässlichen Quellen selbst zu bezahlen, um diese von Hersteller-Interessen unabhängig zu halten.

Eine der großen Chancen von EbM besteht dabei darin, dass wir als Fragende die Definitionshoheit über die Inhalte der Fragen und die verwendeten Quellen zur Beantwortung behalten bzw. zurückerobern. Nur wenn wir immer wieder deutlich machen, welche Fragen sich in der Versorgung von

Patienten tatsächlich stellen, besteht langfristig die Chance, dass sich Anbieter von Information daran orientieren – ein wichtiger Schritt in Richtung des Ziels, Informationsbedarf und Informationsangebot zur Deckung zu bringen.

Tab. 20.1: Anhang – Auswahl nützlicher Quellen und Datenbanken, zugeschnitten auf die allgemeinärztliche Praxis

Quelle (einschließlich Web-Adresse)	Indikation	Erläuterung
Gelbe Liste – www.gelbeliste.de ; Rote Liste – www.rote-liste.de	Frage nach Handels- oder Freinamen von Medikamenten, Darreichungsformen, Packungsgrößen, Dosierung	Die Listen selbst sind kommerzielle Unternehmen; sie verwenden ausschließlich Herstellerangaben; oft in Praxissoftware integriert. Anmeldung erforderlich (Fachkreise)
Fachinformationen – www.fachinfo.de	Wie gelbe und rote Liste, zusätzlich Neben- und Wechselwirkungen, Pharmakokinetik	Beruhend ebenfalls auf Herstellerangaben, aber nach engen Vorgaben der Zulassungsbehörden. Anmeldung erforderlich (Fachkreise)
Arzneimittel-Telegramm – www.arznei-telegramm.de	Frage nach Belegen der Wirksamkeit (Studien evidenz), Nebenwirkungen, Kosten-Nutzen-Verhältnis	Kritische Berichterstattung über Medikamente, möglichst an klinischen Studien orientiert; anzeigenfrei. Erweitertes Angebot für Abonnenten
Infomed-Screen – www.infomed.org	Frage nach Belegen der Wirksamkeit (Studien evidenz)	kritische Berichterstattung neuer Studienerkenntnisse aus hausärztlicher Sicht; anzeigenfrei
Leitlinien der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) – http://leitlinien.net	Frage nach diagnostischem oder therapeutischem Vorgehen, „state of the art“	Heterogene Qualität, da abhängig von einzelner Fachgesellschaft; Qualität anhand Eigenklassifizierung gut erkennbar. Überwiegend Sicht des Spezialisten, aber auch Zugang zu den Leitlinien der Dt. Gesellschaft für Allgemeinmedizin (DEGAM)
Hausärztliche Leitlinien Pharmakotherapie Hessen – www.kvhessen.de	Frage nach therapeutischem Vorgehen bei häufigen Erkrankungen	Beruhet vielfach auf anderen Leitlinien, hausärztliche Sicht. Entstanden im Rahmen der Pharmakotherapieberatung der KV Hessen
Scottish Intercollegiate Guidelines Network (SIGN) – www.sign.ac.uk	Frage nach diagnostischem und therapeutischem Vorgehen bei häufigen Erkrankungen	Weltweit anerkannter Entwicklungsstandard, in englischer Sprache
Clinical Evidence – www.ClinicalEvidence.com	Stand der Evidenz bei häufigen Erkrankungen bzw. Problemen (v.a. Therapie)	S. Kap. 20.4.3

Tab. 20.2: Anhang – Weitere Informationsquellen, die in Kapitel 20 erwähnt wurden**ACP Journal Club**

Zeitschrift, herausgegeben vom American College of Physicians; erscheint alle zwei Monate; enthält zusammengefasste und bewertete Studien primär aus dem Fachgebiet der Inneren Medizin; deckt 200 Fachzeitschriften ab; kostenpflichtig.

Evidence Based Medicine

Zeitschrift, gemeinsam herausgegeben von der McMaster Universität in Ontario (Kanada) und dem Centre for Evidence-Based Practice in Oxford (Großbritannien); erscheint alle 2 Monate; richtet sich primär an Versorgende aus dem hausärztlichen Bereich; enthält zusammengefasste und bewertete Studien aus diesem Fachgebiet; kostenpflichtig.

MEDLINE

Medizinische Literaturdatenbank der amerikanischen National Library of Medicine mit bibliographischen Angaben und Abstracts der wichtigsten wissenschaftlichen Publikationen seit 1966. Enthalten sind die bibliographischen Daten aus ca. 4 000 Zeitschriften aus 70 Ländern zu allen Bereichen der Medizin, größtenteils mit Abstracts. Durch den MESH-Thesaurus (Medical Subject Headings) ist der thematische Zugriff auf die Daten gezielt und umfassend.

Die Datenbasis von MEDLINE entspricht inhaltlich dem (gedruckten) Index Medicus.

PubMed

PubMed ist die kostenfreie Version der Datenbank MEDLINE. Die Datenbank ist zu erreichen über „Entrez (<http://www.ncbi.nlm.nih.gov/Entrez/index.html>)“, das zentrale Suchsystem für die wichtigsten medizinischen Datenbanken.

<http://www.ncbi.nlm.nih.gov/PubMed/>

Im Anhang (s. S. 513ff.) findet sich ein Beispiel einer exemplarischen Suche in MEDLINE über PubMed, in dem die ersten Schritte einer Datenbankrecherche dargestellt sind.

Literatur

- [1] Lorenz KA et al., A qualitative examination of primary care providers' and physicians managers' uses and views of research evidence. *Int J Qual Health Care.* (2005), 17, 409–414
- [2] Curley SP, Connelly DP, Rich EC, Physicians' use of medical knowledge resources: preliminary theoretical framework and findings. *Med Decis Making.* (1990), 10, 231–241
- [3] McColl A et al., General practitioner's perceptions of the route to evidence based medicine: a questionnaire survey. *BMJ.* (1998), 316, 361365
- [4] Slawson DC, Shaughnessy AF, Bennett JH, Becoming a medical information master: feeling good about not knowing everything. *J Fam Pract.* (1994), 38, 505–513
- [5] Ebell MH, Shaughnessy A, Information mastery: integrating continuing medical education with the information needs of clinicians. *J Contin Educ Health Prof.* (2003), 23 Suppl 1, S53–S62
- [6] Putnam W et al., A qualitative study of evidence in primary care: what the practitioners are saying. *CMAJ.* (2002), 166, 1525–1530
- [7] Donner-Banzhoff N et al., Der Evidenzbasierte Praktiker. Ein Beitrag zum hausärztlichen Informationsmanagement. *ZFA.* (2003), 79, 501–509
- [8] Haynes RB, Of studies, syntheses, synopses, and systems: the „4S“ evolution of services for finding current best evidence. *ACP J Club.* (2001), 134, A11–A13
- [9] Garg AX et al., Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA.* (2005), 293, 1223–1238
- [10] Alper BS, White DS, Ge B, Physicians answer more clinical questions and change clinical decisions more often with synthesized evidence: a randomized trial in pri-

- mary care. *Ann Fam Med.* (2005), 3, 507–513
- [11] Pitkin RM, Branagan MA, Burmeister LF, Accuracy of data in abstracts of published research articles. *JAMA.* (1999), 281, 1110–1111
- [12] Gabbay J, le MA, Evidence based guidelines or collectively constructed „mindlines?“ Ethnographic study of knowledge management in primary care. *BMJ.* (2004), 329, 1013
- [13] Andre M et al., Asking for ‚rules of thumb‘: a way to discover tacit knowledge in general practice. *Fam Pract.* (2002), 19, 617–622

III EbM im ärztlichen Alltag

21	Ärztliches Denken und Entscheiden	261
	<i>Norbert Donner-Banzhoff</i>	
22	Evidenzbasierte Medizin im klinischen Alltag – Chirurgie	267
	<i>Stefan Sauerland, Karl-Heinz Moser, Edmund A. M. Neugebauer</i>	
23	Medikamentöse Therapie des Morbus Alzheimer: Ein Disput um die Evidenz zwischen Studienkritik und Klinik	271
	<i>Hanna Kaduszkiewicz, Claus Wächtler</i>	
24	Mammographie-Screening/Prävention – Kritische Bewertung des Screening/Patienteninformation	277
	<i>Ingrid Mühlhauser</i>	
25	EbM in der Rheumatologie – die interdisziplinäre Leitlinie: Management der frühen Rheumatoiden Arthritis	283
	<i>Monika Lelgemann, Matthias Schneider</i>	
26	Evidenzbasierte Medizin in der Allgemeinarztpraxis	289
	<i>Annette Becker, Norbert Donner-Banzhoff</i>	
27	Individualisierte Risikoprognose für die Herz-Kreislauf-Prävention: ein Werkstattbericht	295
	<i>Norbert Donner-Banzhoff, Heidemarie Keller, Eva-Maria Sadowski, Tanja Krones, Erika Baum, Andreas C. Sönnichsen, Uwe Popert</i>	
28a	Revaskularisation bei koronarer Herzerkrankung	305
	<i>Martin Ruß, Karl Werdan</i>	
28b	Chirurgische Revaskularisation bei koronarer Herzkrankheit	313
	<i>Robert von Wattenwyl, Hans-Reinhard Zerkowski</i>	
29	Evidenzbasierte Medizin am Beispiel des lumbalen Bandscheibenvorfalles	319
	<i>Josef Ramsbacher</i>	
30	Schlaganfall und Rehabilitation	325
	<i>Günter Seidel</i>	
31	Evidenzbasierte Medizin in der Pädiatrie – Beispiel Pseudokrupp	331
	<i>Johannes Forster, Dirk Bassler</i>	
32	Evidenzbasierte Medizin bei psychischen Erkrankungen	335
	<i>Michael M. Berner, Mathias Berger, Martin Härter</i>	

33	Evidenzbasierte Medizin in der Sozialmedizin: Beispiel eines Grundsatzgutachtens zur Vertebroplastie und Kyphoplastie	345
	<i>Johannes Giehl, Christoph Kreck</i>	
34	Evidenzbasierte Medizin in der Zahnmedizin – Beispiel Myoarthropathien des Kausystems	351
	<i>Jens C. Türp</i>	

21 Ärztliches Denken und Entscheiden

Norbert Donner-Banzhoff

21.1 Szenario

In der Pause einer Fortbildungsveranstaltung lässt ein Kollege die Bemerkung fallen, die Evidenzbasierte Medizin sei gescheitert. Auf meine Nachfrage hin erläutert er, dass die von Sackett vorgeschlagene Arbeitsweise völlig an der Praxis vorbei gehe. Bei der Tagung ging es um die Wirkungen und Nebenwirkungen der Hormonersatz-Therapie; insgesamt wurden die Teilnehmer mit mindestens zwei Dutzend Forest-Plots (grafische Darstellung von Maßnahmen der Wirksamkeit – Medikamente, Risikofaktoren – einschließlich Konfidenzintervallen, meist im Rahmen einer Metaanalyse) konfrontiert. Die meisten Vortragenden beriefen sich bei ihrer Argumentation ausführlich auf klinische Studien.

Zur Definition der Evidenzbasierten Medizin (EbM) wird immer wieder auf Sacketts klassische „Fünf Schritte“ verwiesen [1]:

- ▲ Frage formulieren,
- ▲ Datenbank-Recherche,
- ▲ kritische Lektüre einer oder mehrerer Originalarbeiten,
- ▲ Formulierung einer Konsequenz,
- ▲ Evaluation.

Dabei stellt das Critical Appraisal einer Originalpublikationen die meisten Ansprüche an Zeit und spezifisches Können.

Legt man diese Definition zu Grunde, ist die EbM tatsächlich gescheitert. Heute wie zur Entstehungszeit der EbM vor 15 Jahren sind Ärztinnen und Ärzte gezwungen, ihre Patienten in einem schnellen Zeittakt zu behandeln. Sie verlassen sich auf Routinen und denken nur über komplizierte und/oder ungewöhnliche Fälle länger nach. Sacketts

„Fünf Schritte“ erscheinen vor diesem Hintergrund als praxisferner Luxus.

Heißt dies, dass die meisten Patienten nicht entsprechend der EbM behandelt werden? Ich möchte hier Zweifel anmelden. Tatsächlich hat die EbM unser Denken und Entscheiden revolutioniert. Dies lässt sich aber nur verstehen, wenn wir von Sacketts „Fünf Schritten“ abstrahieren.¹

21.2 Das kognitive Kontinuum

Ärztliche Überlegungen zu Diagnose und Management lassen sich auf dem so genannten kognitiven Kontinuum einordnen [3]. Dieses reicht von schnellen, automatisch und weitgehend unbewusst ablaufenden Prozessen auf der einen Seite bis hin zu bewussten, überlegten und deshalb langsamen Prozessen auf der anderen Seite. Man nennt die Pole dieses Kontinuums auch „intuitiv“ und „analytisch“; allerdings muss man sich klar machen, dass dieser Intuitionsbegriff vom üblichen Verständnis abweicht [4]. Er beinhaltet überwiegend schlichte, automatisierte Abläufe wie das Gehen oder das Radfahren. In der Medizin wären dies routinierte erfolgreiche Fragen zur Anamnese (z.B. nach Beinödemen bei Dyspnoe) oder Laboranordnungen (z.B. TSH bei Vorhofflimmern). „Analytisch“ gehen wir dagegen bei ungewohnten Fällen vor, die nicht in ein Routineschema passen oder Widersprüche bieten.

¹ Die Überlegungen dieses Kapitels wurden teilweise bereits in folgendem Aufsatz dargestellt: Donner-Banzhoff N. Die Evidenzbasierte Medizin: gescheitert oder einfach nur anders? [2]

Hier denken wir länger nach und konsultieren unter Umständen andere Quellen (z.B. im kollegialen Gespräch). Davon abzugrenzen ist die Eingebung (engl. insight), wenn uns „wie ein Blitz“ die Lösung eines schwierigen Problem in den Sinn kommt, die vielleicht in einer ganz neuen Verknüpfung vertrauter Begriffe besteht [4].

21.3 Der Sinn der Routine

Der analytische Teil des Kontinuums stellt die phylogenetisch jüngere und höher stehende Errungenschaft dar. Die meisten in der Praxis (oder auch im Alltagsleben) auftauchenden Problem lösen wir jedoch mit kaum bewussten Routinen [4]. Der größte Teil der Aufgaben, die sich dem Erfahrenen in seiner Praxis stellen, sind „gewöhnlich“, d.h. für ihre Lösung stehen Routinen zur Verfügung, die mit einem relativ geringen mentalen Aufwand durchgeführt werden können.

Nicht nur die typischen medizinischen Settings („volles Wartezimmer“), sondern ein Grundcharakteristikum unseres Verstandes macht den Einsatz von Routinen unausweichlich: wir haben einen „Arbeitsspeicher“ mit nur sehr begrenzter Kapazität. Zwar verfügt unser Langzeit-Gedächtnis über ein sehr großes Fassungsvermögen; wir haben jedoch Schwierigkeiten, daraus situationgerecht einschlägige Informationen schnell zu mobilisieren und eine größere Menge davon gleichzeitig in unserem Kurzzeit-Gedächtnis zu verarbeiten. Nur wenn wir eine möglichst große Zahl von Aufgaben mit „subkortikalen“ Routinen bewältigen, können wir die zahlreichen Probleme lösen und gleichzeitig einen Teil unseres Arbeitsspeichers den Problemen widmen, die tatsächlich eine bedächtig abwägende Analyse erfordern [5].

21.4 Diagnosen

Untersuchungen zum diagnostischen Vorgehen bei Ärzten kommen übereinstimmend zum Schluss, dass bei einem Patientenkontakt sehr früh fertige kognitive Strukturen („Krankheits-Bilder“) aktiviert werden. Ein induktives Vorgehen, d.h. das sorgfältige Sammeln und Abwägen einer großen Zahl von Informationen (Anamnese, Befund, Labor usw.) findet nur in Ausnahmefällen statt. Die im Gedächtnis vorhandenen Strukturen, die im diagnostischen Prozess aktiviert werden, sind offenbar von früher angetroffenen Patienten „geprägt“. Diese Grundeinheit der ärztlichen Erfahrung hat offenbar die Struktur von Drehbüchern („scripts“), welche narrativ die Informationen zu einem einzelnen Patienten in einen sinnvollen Zusammenhang bringen [6; 7]. Diese früh aufgerufenen Bilder steuern dann die weitere Informationsgewinnung, z.B. im Rahmen der Anamnese. Insgesamt haben wir es also mit einem überwiegend deduktiven Vorgehen zu tun [8].

21.5 Faustregeln und Heuristiken

In vielen anderen Bereichen professionellen Handelns ist gezeigt worden, dass komplizierte und aufwendige Entscheidungsregeln nur selten angewendet werden und im Übrigen auch nicht wirksamer sind als einfache, aber dem jeweiligen Arbeitsmilieu angepasste Heuristiken [5]. In Fokusgruppen schwedischer Allgemeinärzte konnte eine Vielzahl einfacher Faustregeln identifiziert werden, die das alltägliche Versorgungshandeln strukturieren [9].

All diese Dinge müssen bedacht werden, wenn man die Bedeutung der EbM für die Versorgungspraxis erfassen will. Sacketts „Fünf Schritte“ spielen sich an einem Extrem des kognitiven Kontinuums ab, nämlich dem analytischen. Was kann die EbM der

großen Mehrheit der Ärzte bieten, die schon auf Grund äußerer Bedingungen mit halb- bzw. bewussten Routinen und mentalen Short Cuts arbeiten müssen?

21.6 EbM: Instrument der kritischen Reflexion

Zwar bewältigen erfahrene Ärzte den überwiegenden Teil ihrer Arbeit mit Routinen, sie stoßen aber immer wieder an deren Grenzen. Sei es, dass sie mit einem außergewöhnlichen Problem (seltene Symptomatik oder Befund) konfrontiert werden, oder dass das Ergebnis nicht befriedigt oder widersprüchliche Informationen auftauchen. Diese Erfahrungen führen zu „Reflection in action“ [10] oder eben zur Suche nach „externer“ Information, also Evidenz im weitesten Sinne.

Letzteres findet eher selten statt [11]; Zeitmangel, umständlich zu nutzende oder nicht bekannte Datenquellen bzw. die als gering empfundene Aussicht, nützliche Information zu finden, dürften hier eine Rolle spielen. Hier haben die EbM und die mit ihr verbundene technologische Entwicklung (elektronische Datenbanken, Internet) zu einer nachhaltigen Veränderung geführt. Neben den traditionellen Quellen (Lehrbuch, gesammelte Zeitschriftenausschnitte, kollegiales Gespräch) stehen nicht nur quantitativ und qualitativ wesentlich mehr Informationen zur Verfügung als vor 20 Jahren; es gibt außerdem ein differenziertes Rüstzeug, diese Informationen kritisch einzuschätzen. Damit hat die Zuversicht des einzelnen Arztes zugenommen, tatsächlich eine valide Antwort auf seine Frage zu finden. Aus der Psychologie wissen wir, dass diese Zuversicht eine entscheidende Determinante dafür ist, dass ein zielgerichtetes Verhalten tatsächlich erfolgt [12].

Die oben beschriebene narrative Grundstruktur des ärztlichen Wissens wird sich dadurch nicht grundsätzlich ändern. Was wir

aus wissenschaftlichen Studien und den daraus abgeleiteten Quellen erfahren, gibt uns jedoch die Möglichkeit, die in unseren Drehbüchern verwendeten Begriffe und Verknüpfungen kritisch zu reflektieren. So hilft uns eine Untersuchung zur Diagnose bei der Entscheidung, ob ein bestimmter Test überhaupt Teil unserer patientenbezogenen Drehbücher sein soll. Auch unsere Annahme über den Zusammenhang von Risikofaktoren und Erkrankungen, die sich wiederum in unseren Drehbüchern niederschlägt, können wir mithilfe entsprechender Studien hinterfragen. Bei der Entscheidung darüber, ob wir mit therapeutischem Handeln in das Drehbuch eingreifen sollen, können klinische Studien mit patientenrelevanten Zielkriterien eine wichtige Hilfe sein.

Dabei hatte schon die McMaster-Gruppe ein breites Spektrum im Sinn, mit dem die EbM das alltägliche Handeln informieren sollte. Sie definierten den „Appraising Mode“ (die klassischen „Fünf Schritte“), den „Searching Mode“ (schnelle Recherche in evidenzbasierten Quellen) und den konventionellen „Replicating Mode“, in dem wir uns auf Autoritäten bzw. Evidenz-Aufbereitungen verlassen, damit allerdings auch deren Bias ausgeliefert sind [1]. Hier ergibt sich eine Problemlösungs-Pyramide, bei der die meisten Ärzte sich für die Mehrheit ihrer Fragen in der breiten Basis des Replicating und Searching Mode bewegen. Sie benutzen niedrigschwellige, weitgehend aufbereitete Quellen, um Antworten auf konkrete Fragen zu erhalten. Nur wenige Kollegen haben die Kompetenz, den Appraising Mode anhand einer wissenschaftlichen Originalarbeit zu praktizieren; selbst diejenigen, die dazu in der Lage sind, beschreiten aus Zeitgründen diesen Weg nur für wenige ausgewählte Fragen.

21.7 EbM-Rezeption: anders als erwartet

Will man den Impact von EbM erfassen, darf man dies nicht an der Zahl der Ärzte festmachen, die sich in schwer verständliche, englischsprachige Originalarbeiten vertiefen. Ich muss allerdings zugeben, dass ich hier mangels empirischer Daten weitgehend auf der Basis persönlicher Eindrücke argumentiere. Zu Beginn der EbM war noch nicht abzusehen, welche Veränderungen sich ergeben würden, sodass vergleichende Untersuchungen (vorher – nachher) des ärztlichen Reflektierens, Entscheidens und Handelns nicht prospektiv definiert werden konnten.

Zum Pflichtstoff von Pharmakotherapie-Qualitätszirkeln gehören heute ganz selbstverständlich Begriffe wie die Absolute und Relative Risiko-Reduktion oder die Number needed to treat. Leitlinien sind zur allgemein akzeptierten Hilfe ärztlichen Handelns geworden. Allmählich wächst auch das Bewusstsein dafür, dass hier der Konsens von Autoritäten nicht genügt, sondern eine Absicherung durch Studienevidenz erforderlich ist. Wie das zu Beginn referierte Szenario zeigt, ist die kollegiale Diskussion heute in viel stärkerem Maß vom Bezug auf klinische Studien geprägt als vor zehn Jahren; vielfältige Aus-, Weiter- und Fortbildungsmöglichkeiten vermitteln die Fähigkeit, solche Studien kritisch zu rezipieren. Dies geht mit einer positiven Einstellung zur EbM einher, beispielsweise bei niedergelassenen Ärztinnen und Ärzten [13]. Studien, welche die Evidenzbasierung in Praxis und Klinik untersuchen, haben schon in der Frühzeit der EbM-Bewegung einen hohen Anteil entsprechend fundierter Maßnahmen aufgezeigt [14; 15].

21.8 EbM – nicht jedem willkommen

Die Ärzteschaft zeigt das ganz Spektrum zwischen Begeisterung für EbM und tiefer Abneigung dagegen. Eine eher skeptische Einstellung gegenüber meinungsbildenden Autoritäten und Traditionen, ein aktives Verhältnis zur eigenen Fortbildung und die Bereitschaft, Qualitätsmanagement i.w.S. produktiv zu nutzen, sind Eigenschaften, die besonders „anfällig“ für die EbM machen. Diese innovativ gesonnene Gruppe hat jetzt ein wirksames Instrumentarium in der Hand, Antworten auf ihre Fragen zu finden. Dies wird eher selten in der Form von Sackets „Fünf Schritten“ geschehen; überwiegend werden auch hier aufbereitete Datenquellen, Heuristiken und Faustregeln [16] zur Anwendung kommen. Trotzdem entspricht dies dem Kern von EbM: den Zweifel akzeptieren, daraus Fragen formulieren, ernst nehmen, Antworten suchen und die Ergebnisse kritisch diskutieren (s. Kap. 1). Genau dies war bei der McMaster-Gruppe der Ausgangspunkt der EbM [17].

Allerdings kommen an diesem Punkt doch wieder die „Fünf-Schritte“ ins Spiel, nämlich dann, wenn man sich Gedanken darüber macht, welche Kompetenzen zur Verfolgung einer Frage erforderlich sind. Die „Fünf Schritte“ umschreiben zwar eine nur ganz selten geübte **Praxis**, andererseits aber doch ein essentielles Spektrum von ärztlichen **Kompetenzen** im Informationszeitalter. Wir sollten also die „Fünf Schritte“ als ein **edukatives Paradigma** sehen, das aber nur zusammen mit dem Hinweis auf die vielen Formen vermittelt werden sollte, in denen die Evidenzbasierte Medizin lebendig geworden ist und sich weiterentwickeln wird.

Literatur

- [1] Sackett DL, Straus SE, et al. (2000) Evidence-Based medicine. How to practice and teach EbM. Churchill-Livingstone, New York
- [2] Donner-Banzhoff N, Die Evidenzbasierte Medizin: gescheitert oder einfach nur anders? Z Arztl Fortbild Qualitätssich. (2006), im Druck
- [3] Hamm RM (1988) Clinical intuition and clinical analysis: Expertise and the cognitive continuum. In: Dowie J, Elstein A, Professional judgement. A reader in clinical decision making, 78–105. Cambridge Univ. Pr., Cambridge
- [4] Hogarth RH (2001) Educating intuition. Univ. Chicago, Chicago
- [5] Gigerenzer G, Todd PM, The ABC research group (1999) Simple heuristics that make us smart. Oxford Univ. Pr., Oxford
- [6] Tanenbaum SJ, Knowing and acting in medical practice: the epistemological politics of outcomes research. J Health Polit Policy Law. (1994), 19, 27–44
- [7] Schmidt HG, Norman GR, Boshuizen HP, A cognitive perspective on medical expertise: theory and implication. Acad Med. (1990), 65, 611–621
- [8] Elstein AS, Shulman LS, Sprafka SA (1978) Medical problem-solving: An analysis of clinical reasoning. Harvard Univ. Pr., Cambridge
- [9] Andre M et al., Asking for ‚rules of thumb‘: a way to discover tacit knowledge in general practice. Fam Pract. (2002), 19, 617–622
- [10] Schön DA (1987) Educating the reflective practitioner. Jossey-Bass, San Francisco
- [11] Ely JW et al., Answering physicians' clinical questions: obstacles and potential solutions. J Am Med Inform Assoc. (2005), 12, 217–224
- [12] Bandura A, Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev. (1977), 84, 191–215
- [13] Meyer T, Ströbel A, Raspe H, Evidenzbasierte Medizin aus Sicht niedergelassener Ärztinnen und Ärzte: ein repräsentativer Survey zu Akzeptanz und Fortbildungsinteressen. Z Arztl Fortbild Qualitätssich. (2004), 98, 293–300
- [14] Gill P et al., Evidence based general practice: a retrospective study of interventions in one training practice. BMJ. (1996), 312, 819–821
- [15] Ellis J et al., Inpatient general medicine is evidence based. A-Team, Nuffield Department of Clinical Medicine. Lancet. (1995), 346, 407–410
- [16] Donner-Banzhoff N et al., Der Evidenzbasierte Praktiker. Ein Beitrag zum hausärztlichen Informationsmanagement. ZFA. (2003), 79, 501–509
- [17] Sackett DL, How to read clinical journals: 1. Why to read them and how to start reading them critically. CMAJ. (1981), 124, 555–590

22 Evidenzbasierte Medizin im klinischen Alltag – Chirurgie

Stefan Sauerland, Karl-Heinz Moser, Edmund A. M. Neugebauer

22.1 Klinisches Szenario

Ein 25-jähriger Student kommt humpelnd und durch eine Freundin gestützt in das Sprechzimmer Ihrer chirurgischen Praxis. Er erklärt Ihnen, er sei am Vorabend beim Volleyball-Spielen nach einem Blocksprung mit seinem rechten Fuß auf dem Fuß eines Mitspielers gelandet und umgeknickt. Er habe unmittelbar nach der Verletzung das Spiel abgebrochen und den Knöchel mit Eis gekühlt. In der Nacht habe er das Bein erhöht gelagert, es sei aber dennoch schnell zu einer erheblichen Schwellung gekommen. Der Patient berichtet Ihnen auf Nachfragen, dass er ohne Hilfsmittel das Spielfeld verlassen konnte und das betroffene Bein dabei auch voll belastete. Frühere Verletzungen oder Vorerkrankungen sind nicht bekannt.

Bei der Untersuchung findet sich auf der Außenseite des Knöchels eine erhebliche, teils blutig unterlaufene Schwellung. Aktive und passive Beweglichkeit, Durchblutung und sensible Innervation sind unauffällig. Im Bereich des fibulo-talaren Übergangs lässt sich ein umschriebener Druckschmerz auslösen. Der Einbeinstand auf dem betroffenen Fuß ist dem Patienten unter Schmerzen möglich. Auch der Fersenstand und Hackenstand gelingen. Sie überlegen, welche weitere Diagnostik und Therapie indiziert ist.

werden, entsprechend den vier Freiheitsgraden der Bewegung im oberen (OSG) und unteren Sprunggelenk (USG). Die Supinations-Eversionsverletzung ist mit 40–70% die häufigste. Dabei sind sowohl knöcherne als auch ligamentäre Verletzungen entlang der Supinator-Linie nach Hellpap möglich [1]. Neben Außenbandrupturen (Ein-, Zwei- oder Drei-bandrupturen) und Syndesmosenverletzungen sind bei diesem Verletzungsmuster auch Fibulafrakturen, Abrissfrakturen an der Basis des Metatarsale V und knöcherne Verletzungen der Tibia (knöcherner Abriss der Syndesmose) in ca. 20% der Fälle zu beobachten.

Schmerzen, Schwellung, Blutergüsse und eine Funktionseinschränkung sind sowohl bei Bandverletzungen als auch bei Frakturen nachweisbar und deshalb unspezifisch. Während man beim Supinations-Eversionstrauma unter entsprechenden Voraussetzungen auf Röntgenaufnahmen verzichten kann, ist bei der Supinations-Adduktions-, Pronations-Abduktions- und Pronations-Eversionsverletzung fast immer mit einer Fraktur zu rechnen, sodass die klinische Untersuchung in diesen Fällen mit einer Röntgenaufnahme in 2 Ebenen ergänzt werden muss und – bei entsprechendem Verdacht – noch zusätzliche Zielaufnahmen erforderlich sein können.

22.2 Medizinischer Hintergrund

Verletzungen des oberen Sprunggelenks treten jährlich bei 50–60 von 10000 Einwohnern auf. Die Sprunggelenksdistorsion ist die häufigste Verletzung. Es können vier verschiedene Verletzungstypen unterschieden

22.3 Die Evidenz

Die Ottawa Ankle Rule (OAR) ist ein inzwischen weitverbreitetes Instrument, um bei Verletzungen des Knöchels oder Rückfußes über die Notwendigkeit radiologischer Diagnostik entscheiden zu können. Eine Arbeits-

gruppe um den kanadischen Arzt Ian G. Stiell hat die OAR Anfang der 90er-Jahre zuerst entwickelt, dann verfeinert und schließlich ausgiebig validiert [2–4]. Eine akut behandlungsbedürftige Fraktur kann nach der OAR bei jedem Patienten ausgeschlossen werden, der nach dem Unfall gelaufen ist, bei der Untersuchung vier Schritte ohne Hilfsmittel gehen kann und bei Druck auf die Hinterkante des Außenknöchels, Druck auf die Basis des Metatarsale V, Druck auf die Hinterkante des Innenknöchels und Druck auf das Os naviculare keinen wesentlichen Schmerz angibt (vgl. Abb. 22.1). Laut der OAR ist umgekehrt eine Röntgenuntersu-

chung des Knöchels nur dann indiziert, wenn sich über der Hinterkante des medialen oder lateralen Malleolus ein Druckschmerz auslösen lässt oder der Patient das Bein nicht voll belasten kann. Eine radiologische Untersuchung des Fußes ist nur notwendig, wenn ein Druckschmerz über der Basis des fünften Os metatarsale oder dem Os naviculare besteht oder wiederum eine Vollbelastung des Fußes unmöglich ist.

Inzwischen haben über 30 Studien die Ottawa Ankle Rule klinisch getestet. Eine Meta-Analyse dieser diagnostischen Studien [5] konnte aus den Ergebnisse von über 10000 Patienten die folgenden Zahlen er-

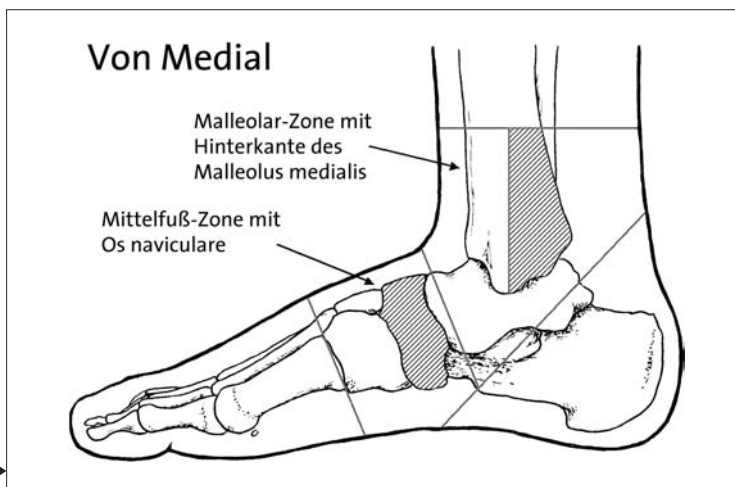
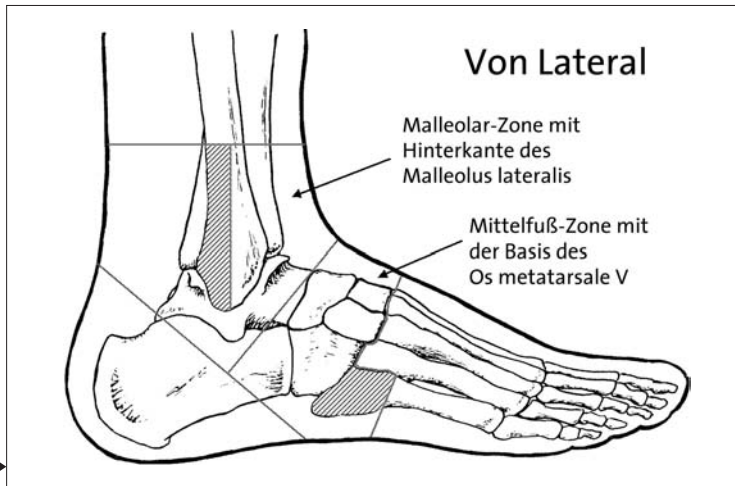


Abb. 22.1a, b: Die Ottawa Ankle Rule: 1. Röntgendiagnostik ist bei Schmerzen in der Malleolar-Zone nur erforderlich, wenn zusätzlich ein Druckschmerz über einer der beiden Malleolen-Hinterkanten (6 cm) vorliegt oder der Patient seit dem Unfall den Fuß nicht wieder belasten konnte. 2. Röntgendiagnostik ist bei Schmerzen in der Mittelfuß-Zone nur erforderlich, wenn zusätzlich ein Druckschmerz über der Basis des Os metatarsale V oder dem Os naviculare auslösbar ist oder der Patient seit dem Unfall den Fuß nicht wieder belasten konnte.

mitteln: Die OAR besitzt eine Sensitivität von etwa 98% und eine Spezifität von 40%. Die negative Likelihood-Ratio liegt bei 0,08. Dieser extrem niedrige Wert bedeutet, dass ein unauffälliger Befund gemäß der Ottawa Ankle Rule das Vorliegen einer Fraktur so gut wie ausschließen kann. Selbst wenn man primär von einer Grundwahrscheinlichkeit (also Prävalenz) von 15% ausgeht, beträgt das rechnerische Restrisiko, mit der OAR eine Fraktur zu übersehen, nur 1% [5].

22.4 Auflösung des Szenarios

Da im vorliegenden Fall bei dem geschilderten Supinations-Eversionstrauma keines der diagnostischen Kriterien der Ottawa Ankle Rules positiv ist, kann man eine dislozierte oder operationspflichtige Knöchelfraktur mit hinreichender Sicherheit ausschließen. Auf eine radiologische Diagnostik zum Frakturausschluss kann also verzichtet werden.

Als mögliche Diagnose verbleibt neben einer harmlosen Distorsion eine Außenbandzerreiung. Bei der Diagnostik der ligamentären Schden gibt es eine Vielzahl klinischer Tests, die eine relativ hohe Validitt haben. In der klinischen Praxis hat die Bestimmung des Talusvorschubs eine hinreichend hohe diagnostische Genauigkeit. Dabei liegt der Patient mit dem Unterschenkel auf der Untersuchungsliege, der Fuß hngt frei. Der Untersuchende fixiert den Unterschenkel mit der einen Hand und bt mit der anderen Hand einen dosierten Druck von hinten auf das Fersenbein aus, sodass es zu einer ventralen Verschiebung im Sprunggelenk kommt. Da die Bnder von Individuum zu Individuum unterschiedlich lax sein knnen, gilt erst ein vermehrter Talusvorschub von >4 mm im Seitenvergleich als Beweis fr eine Auenbandruptur.

Ausgehend von dem anamnestisch beschriebenen Supinations-Eversionstrauma sollten Sie je nach Ergebnis Ihres Untersu-

chungsbefund deshalb folgende Therapie einleiten:

- ▲ negative OAR + negativer Talusvorschub: Hochlagerung, Kryotherapie, nicht-steroidale Antiphlogistika, Tape oder elastischer Kompressionsverband,
- ▲ negative OAR + positiver Talusvorschub: Knchelorthese, Kryotherapie, nicht-steroidale Antiphlogistika, Peroneusstimulation,
- ▲ positive OAR: Primre Rntgenuntersuchung und Weiterleitung zum Facharzt.

Es ist sicherzustellen, dass diese Patienten engmaschig kontrolliert werden und sptestens nach 5 Tagen nochmals bei einem Facharzt vorgestellt werden. Im Zweifelsfall oder bei Beschwerdepersistenz muss doch die bildgebende Diagnostik durchgefhrt werden, um keine Schden am OSG bzw. USG zu bersehen, die spter die Ursache chronischer Schmerzen sein knnten.

22.5 Kommentierung und Ausblick

Die Ottawa Ankle Rule ist eine der am besten berprften klinischen Entscheidungsregeln, die es in der Medizin gibt [5]. Die klinischen und konomischen Vorteile der OAR resultieren direkt aus der Reduktion der Rntgendiagnostik um etwa 25%, wie sie in mehreren Studien beobachtet wurde [5]. Auch wenn die OAR primr fr Erwachsene entwickelt wurde, wird sie inzwischen auch bei Jugendlichen und sogar Kindern angewendet. Ferner gab es erfolgreiche Versuche, nichtrztliches Personal mit der Anwendung der Ottawa Ankle Rule zu betrauen [6]. In Kanada hngt in 60% der Notaufnahmen die OAR als Poster aus [7].

Auch wenn die Ottawa Ankle Rule sich weltweit verbreitet und die Versorgung beeinflusst hat [8–10], so mssen doch auch gewisse Probleme mitbetrachtet werden. Lngst nicht alle rzte, die die OAR verwen-

den, kennen sie wirklich genau und wenden sie exakt an [7; 11]. So ergab eine kanadische Umfrage, dass mehr als die Hälfte der Ärzte die Ottawa Ankle Rule dann nicht anwenden, sondern ein Röntgenbild erstellen lassen, wenn eine starke Schwellung am Außen- oder Innenknöchel vorliegt, wenn altersbedingt ein osteoporotisch erhöhtes Frakturrisiko vorliegt oder wenn ein Sturz aus größerer Höhe die Unfallursache war [7]. Solche Variationen einer validierten Entscheidungsregel sind aus medizinisch-epidemiologischer Sicht logisch nicht begründbar, sondern deuten auf ein tief verwurzeltes Sicherheitsdenken hin, das nur langsam verdrängt werden kann.

Auch in Deutschland ist die Ottawa Ankle Rule sehr bekannt [8], wird offenbar aber immer noch nicht in allen Kliniken angewendet. Es muss in diesem Zusammenhang bedacht werden, dass die Angst vor haftungsrechtlichen Folgen einer unvollständigen Diagnostik weiterhin sehr groß ist. Auch werden bei berufsgenossenschaftlich zu bewertenden Fällen aus Gründen der juristisch-gutachterlichen Dokumentation vielfach Röntgenbilder benötigt. Schließlich muss bedacht werden, dass die Entscheidung, auf ein Röntgenbild zu verzichten, auch mit dem Patienten klar besprochen werden muss. In etwa einem Viertel der Fälle erwartet der Patient, dass eine radiologische Abklärung erfolgt. Hier kann nur ein offenes Arzt-Patienten-Gespräch helfen, bei dem auch auf die klare Evidenz hingewiesen wird.

Aufgrund des enormen klinischen Nutzens der OAR entstanden in der Folgezeit in Kanada weitere ähnlich entwickelte Regeln, um nach Verletzungen der Kniegelenks, des Schädels oder der Halswirbelsäule über die Notwendigkeit radiologischer Untersuchungen entscheiden zu können. Auch diese Entscheidungsregeln haben sich teils aus medizinischen, teils aus ökonomischen Gründen auf breiter Front durchsetzen können und haben entscheidend dazu beigetragen, die

Strahlenbelastung chirurgischer Patienten zu reduzieren.

Literatur

- [1] Hellpap W, Das vernachlässigte untere Sprunggelenk. Die „Frakturlinie“ der Supination. Arch Orthop Unfallchir. (1963), 55, 289–300
- [2] Stiell IG et al., Decision rules for the use of radiography in acute ankle injuries. Refinement and prospective validation. JAMA. (1993), 269, 1127–1132
- [3] Stiell I et al., Multicentre trial to introduce the Ottawa ankle rules for use of radiography in acute ankle injuries. Multicentre Ankle Rule Study Group. BMJ. (1995), 311, 594–597
- [4] Stiell IG et al., A study to develop clinical decision rules for the use of radiography in acute ankle injuries. Ann Emerg Med. (1992), 21, 384–390
- [5] Bachmann LM et al., Accuracy of Ottawa ankle rules to exclude fractures of the ankle and mid-foot: systematic review. BMJ. (2003), 326, 417
- [6] Fiesseler F et al., Can nurses appropriately interpret the Ottawa Ankle Rule? Am J Emerg Med. (2004), 22, 145–148
- [7] Brehaut JC et al., Clinical decision rules „in the real world“: how a widely disseminated rule is used in everyday practice. Acad Emerg Med. (2005), 12, 948–956
- [8] Chandra A, Schafmayer A, Die diagnostische Wertigkeit eines klinischen Tests zum Ausschluss von Frakturen nach akuten Sprunggelenkdorsionen. Eine prospektive Studie zur Überprüfung der „Ottawa Ankle Rules“ in Deutschland. Unfallchirurg. (2001), 104, 617–621
- [9] Graham ID et al., Awareness and use of the Ottawa ankle and knee rules in 5 countries: can publication alone be enough to change practice? Ann Emerg Med. (2001), 37, 259–266
- [10] Holroyd BR et al., Uptake of validated clinical practice guidelines: experience with implementing the Ottawa Ankle Rules. Am J Emerg Med. (2004), 22, 149–155
- [11] Graham ID et al., Emergency physicians' attitudes toward and use of clinical decision rules for radiography. Acad Emerg Med. (1998), 5, 134–140

23 Medikamentöse Therapie des Morbus Alzheimer: Ein Disput um die Evidenz zwischen Studienkritik und Klinik

23.1 Hintergrund

Hanna Kaduszkiewicz, Claus Wächtler

Über den Einsatz von Acetyl-Cholinesterasehemmern (AChEH) zur Behandlung der Alzheimer-Demenz wird intensiv gestritten. Auf der einen Seite ist das wissenschaftliche und berufspolitische Schrifttum weitgehend einig darüber, dass die AChEH Donepezil, Rivastigmin und Galantamin bei leichter bis mittelschwerer Alzheimer-Demenz die Präparate der ersten Wahl sind. Diese Empfehlung findet sich mit geringen Varianzen in allen nationalen und internationalen Leitlinien [1]. Auf der anderen Seite kommt eine kritische Diskussion über die methodische Qualität der zugrunde liegenden Studien zu dem Ergebnis, dass aufgrund methodischer Mängel dieser Studien der wissenschaftliche Nachweis der Wirksamkeit von AChEH bisher noch nicht erbracht wurde [2]. Des Weiteren wird sowohl die klinische Relevanz der statistisch signifikanten Studienergebnisse infrage gestellt, als auch die Kosteneffektivität der AChEH [3]. Ziel dieses Kapitels ist es, Argumente für und gegen die Verschreibung von AChEH darzustellen.

23.2 Zwei Patienten mit Alzheimer-Demenz

Claus Wächtler

Szenario 1

Seit einem Jahr zeigt Frau D., 69 Jahre alt, zunehmend Gedächtnisstörungen. Sie vergisst sehr

schnell und muss sich mittlerweile alles aufschreiben. Sie verpasst wichtige Termine wie z.B. den Gang zum Arzt. Seit kurzem fallen ihr Worte im Alltag nicht mehr ein. Beim Autofahren weiß sie manchmal nicht mehr, wo sie ist. Mehrmals wusste sie mit dem Autoschlüssel nichts anzufangen. Die Bewältigung ihres Haushaltes fällt ihr zunehmend schwer.

Bei der Untersuchung fallen bei der primär differenzierten Patientin ein einfacher Sprachgebrauch und Wortfindungsstörungen auf. Die weitere psychopathologische Untersuchung und eine Reihe zusätzlich eingesetzter Hirnleistungs-Screening-Tests (z.B. Mini-Mental-State-Examination (MMSE): 20 Punkte) erhärten den Verdacht auf eine beginnende Demenz. Die typische Anamnese und der Ausschluss anderer Erkrankungen machen eine Demenz vom Alzheimerstyp wahrscheinlich.

Szenario 2

Herr B., 73 Jahre alt, kommt wegen einer seit zwei Wochen zunehmenden, mittlerweile schweren Unruhe mit auch aggressiven Äußerungen und zum Teil Tätlichkeiten gegenüber der Familie zur Aufnahme in eine gerontopsychiatrische Klinik. Wegen der Gespanntheit, der Krankheitsuneinsichtigkeit und der Weglauftendenz muss er gemäß PsychKG geschlossen untergebracht werden. Es fällt auf, dass Herr B. teils abwesend wirkt und sich mit Gegenständen oder Personen zu befassen scheint, die nicht anwesend sind. Er ist von ständiger Unruhe befallen.

Von den Angehörigen erfahren wir, dass vor einem Jahr eine Demenz vom Alzheimerstyp diagnostiziert worden war. Auf eine antidementive Therapie mit einem AChEH habe der Patient mit stärkerer Übelkeit, fraglich auch mit Unruhe,

reagiert, sodass das Medikament abgesetzt wurde.

→ Was würden Sie tun? Und warum?

23.3 Der Abwägungsprozess

23.3.1 Methodenkritische Überlegungen

Hanna Kaduszkiewicz

Wer sich auf den Spuren der Evidenzbasierten Medizin bewegt, sollte die Studien am besten persönlich lesen. Ende 2005 waren 12 randomisierte, kontrollierte, doppelblinde Studien (RCT) zu Donepezil und jeweils 5 zu Rivastigmin und Galantamin publiziert. In drei Studien zu Donepezil erreichte keiner der primären Endpunkte die Grenze der statistischen Signifikanz, in allen anderen Studien präsentierten die Autoren einen oder mehrere statistisch signifikante Unterschiede zwischen der Behandlungs- und Placebogruppe – immer zugunsten des AChEH.

Ohne Frage waren alle 22 Studien vom Ansatz her korrekt angelegt. Doch bei Durchsicht der Publikationen zeigten sich in jeder der Studien mehrere methodische Probleme, die einen entscheidenden Einfluss auf das Ergebnis gehabt haben können. Beispiele:

- ▲ Verblindung der Patienten und Untersucher ist ein herausragendes Qualitätskriterium von Studien, weil dadurch der einseitige Einfluss von Wirksamkeitsannahmen unterbunden wird. Dabei müssen die Untersucher auch gegenüber Nebenwirkungen verblindet werden, da sonst die Gruppenzugehörigkeit bei für den Wirkstoff typischen Nebenwirkungen leicht zu erraten ist. Hierzu gibt es in der Mehrzahl der Studien keine Angaben.
- ▲ Viele Patienten (etwa 15% und mehr pro Gruppe) sind vorzeitig aus den Studien ausgeschieden. Um deren fehlende Werte zu ersetzen, wurde die LOCF-Methode

(Last observation carried forward) verwendet. Dabei wird die letzte verfügbare Messung des vorzeitig ausgeschiedenen Patienten als Messung am Studienende definiert. Dieses Verfahren hat bei einer chronisch progredienten Erkrankung wie der Alzheimer-Demenz den Nachteil, dass Gruppen, aus denen viele Patienten ausscheiden, artifiziell ein besseres Ergebnis erzielen.

- ▲ Intention to treat (ITT) bedeutet, dass alle randomisierten Patienten bei den Endpunktberechnungen berücksichtigt werden. Dies ist – trotz des häufig selbst gestellten Anspruchs – in vielen Studien nicht der Fall.

Eine genaue Übersicht über die 22 Studien, ihre Ergebnisse und methodischen Mängel finden Sie in der Publikation von Kaduszkiewicz et al. [2], die im Internet frei zugänglich ist.

Bis Mai 2006 wurden zwei weitere RCTs zu Donepezil veröffentlicht [4; 5]. Die Studie von Rockwood et al. war mit 130 Teilnehmern recht klein. Bei zwei a priori definierten Endpunkten zeigte einer ein statistisch signifikantes Ergebnis zugunsten von Donepezil, der andere keinen Unterschied zwischen den Gruppen. Eine Interpretation des Ergebnisses dieser Studie ist daher schwierig [6]. Die zweite im Jahre 2006 veröffentlichte Studie untersuchte erstmalig die Wirkung von Donepezil bei Patienten mit schwerer Alzheimer-Demenz. Für diese Indikation sind die AChEH bisher nicht zugelassen. Die Autoren scheinen aus den methodenkritischen Diskussionen gelernt zu haben, denn sie präsentieren z.B. eine ITT-Analyse, bei der die Werte der ausgeschiedenen Patienten durch den Mittelwert der Placebogruppe ersetzt werden. Dies ist eine gute Ersatzmöglichkeit – angesichts von 26% ausgeschiedenen Patienten in der Donepezil- und 18% in der Placebogruppe ist die Zulässigkeit eines Ersetzens von Werten in dieser Studie aber

generell infrage zu stellen. Zudem bleiben auch in dieser Studie Fragen offen, z.B. ist die Verblindung nicht beschrieben, es fehlen die Baseline-Testergebnisse für die ITT-Population – und wie bei Rockwood et al. ist von zwei a priori definierten Endpunkten nach Korrektur für multiples Testen ein Endpunkt statistisch signifikant, der andere nicht.

23.3.2 Reviews

Hanna Kaduszkiewicz

Die Autoren der beschriebenen RCTs sehen in der Mehrzahl keine methodischen Probleme in ihren Studien. Ebenso verhält es sich mit den Reviews der Cochrane Collaboration [7–9], die für den Einsatz aller drei Cholinesterasehemmer eine ausreichende wissenschaftliche Evidenz feststellen und größere methodische Probleme nicht sehen. Birks und Harvey zum Beispiel beschreiben bezüglich der Donepezil-Studien Dropout-Raten von unter 20% als „klein“ und bezweifeln eine negative Konsequenz für die Ergebnisse der Studien. Aus biometrischer Sicht ist das jedoch problematisch [10].

23.3.3 Überlegungen zur Übertragbarkeit der Ergebnisse auf „echte Patienten“

Hanna Kaduszkiewicz

Über die Betrachtung der Studienmethodik hinaus müssen zur Begründung eines Einsatzes bzw. Nicht-Einsatzes der AChEH eine Reihe von Fragen zur Übertragbarkeit der Studienergebnisse auf „echte Patienten“ beantwortet werden:

- ▲ Sind statistisch signifikante Mittelwertsunterschiede von 1.5–3.9 Punkten auf einer 70-stufigen kognitiven Skala (ADAS-cog) eigentlich klinisch relevant?

- ▲ Dieselbe Frage gilt für die zur Beurteilung des Gesamtzustandes der Patienten in den Studien häufig verwendete 7-stufige CIBIC-plus-Skala: Was sagt ein Mittelwertunterschied von max. 0.54 Punkten aus?
- ▲ Wie lange sollen Patienten behandelt werden, wenn in 21 der 24 RCTs die Studiendauer 26 Wochen nicht überschritt und zwei der drei Studien, die länger dauerten, keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen gezeigt haben?
- ▲ Für wen gelten die Studienergebnisse, wenn Patienten mit Insulin-abhängigem Diabetes mellitus, anderen endokrinologischen Störungen, Asthma, obstruktiven Lungenerkrankungen oder klinisch signifikanten, nicht kontrollierten kardiovaskulären, gastrointestinalen oder hepatischen Erkrankungen in die meisten Studien nicht aufgenommen wurden?
- ▲ Unbestritten von den Befürwortern der AChEH ist, dass nur ein Teil der Demenzpatienten auf eine Therapie mit AChEH anspricht und dass Responder vorab nicht identifiziert werden können. Ist dies Grund genug, um bei allen Demenzpatienten eine Therapie zu beginnen?
- ▲ Schließlich muss sich der Leser auch die Frage stellen, ob die AChEH eventuell ihre Wirkung nur auf Ebenen entfalten, die mit den üblicherweise in Studien eingesetzten Testinstrumenten nicht erkannt werden können, z.B. weil ihre Skalen zu grob sind.

23.3.4 Erfahrungen von Klinikern und Patienten

Claus Wächtler

Neben der Evidenz aus wissenschaftlichen Untersuchungen spielen im Abwägungsprozess für bzw. gegen den Einsatz eines Medikamentes auch die klinische Erfahrung des Arz-

tes sowie der Wunsch des Patienten bzw. seiner Angehörigen eine wichtige Rolle. Bei der Alzheimer-Demenz handelt es sich um eine Erkrankung mit erheblicher Progression und derzeit noch geringen therapeutischen Optionen. Umso mehr fordern Praktiker und Angehörige (unterstützt durch die Deutsche Alzheimer-Gesellschaft und Alzheimer Europe), dass alle heute verfügbaren Maßnahmen, die aufgrund vorliegender Evidenz oder der Beobachtung in der täglichen Praxis Wirksamkeit versprechen, konsequent eingesetzt werden. Hierzu gehören aus Sicht des Klinikern neben dem gezielten Einsatz psychosozialer Maßnahmen wie kognitiver Anregung gemeinsam mit körperlicher Aktivierung sowie Stützung und Schulung der Angehörigen insbesondere auch der wohl abgewogene und kontrollierte Einsatz von Acetyl-Cholinesterasehemmern. In der Praxis erweist sich aus Sicht des Klinikern gerade der kombinierte Einsatz der heute verfügbaren therapeutischen Maßnahmen als am besten wirksam. Die klinische Erfahrung mit Cholinesterasehemmern zeigt, dass etwa ein Drittel der Patienten deutlich und ein weiteres Drittel mäßig (in der Regel im Sinne einer vorübergehenden Stabilisierung) reagiert. Diese Rückmeldung bestätigen Angehörige sowie Pflegekräfte in ambulanten und stationären Einrichtungen – und, sofern der Einsatz in einem relativ frühen Stadium der Erkrankung erfolgte, auch die Patienten selbst. Unabhängig davon ist zu fordern, dass – neben der Bereitstellung ausreichender Schulung des Personals in psychosozialen Maßnahmen und entsprechender personeller Kapazität – noch wirksamere Medikamente entwickelt werden.

23.3.5 Schließen der Patientenfälle

Claus Wächtler

Um auf die beiden eingangs geschilderten Patienten zurückzukommen: → Was würden Sie jetzt tun? Und warum?

Szenario 1

Bei der Patientin wurde eine Therapie mit einem AChEH begonnen. Bei Verlaufskontrollen nach 3 und 6 Monaten berichtet die Patientin, dass sie sich im Alltag wieder besser zurechtfinde. Es falle ihr leichter, sich Namen und Termine zu merken, beim Autofahren bestünden derzeit keine Probleme. Es falle ihr aber weiterhin schwer, komplizierte Worte zu finden. Die Angehörigen bestätigen das von der Patientin Beobachtete. Bei den Hirnleistungsuntersuchungen findet sich eine leichte Besserung in fast allen Bereichen. Im MMSE erreicht die Patientin nach drei Monaten 23 und nach sechs Monaten 24 Punkte.

Nach einem Jahr wird vom Hausarzt der Cholinesterasehemmer abgesetzt. Er führt als Begründung an, dass er lediglich eine Stabilisierung des Befindens gefunden habe, nicht aber eine Besserung. Außerdem nennt er Budgetgründe. Die Angehörigen berichten, dass nach Absetzen des AChEH eine gravierende Verschlechterung der Demenz eingetreten sei. Die Fähigkeiten zur Alltagsbewältigung hätten binnen kurzem stark abgenommen. Die Patientin könne sich mittlerweile nicht mehr selbst waschen und anziehen. Sie zeige vermehrt Angst, Depression und Unruhe. Daraufhin erhält die Patientin in der Tagesklinik erneut den AChEH sowie das Antidepressivum Citalopram, Gruppentherapie und Ergotherapie. Darunter zeigt sich eine deutliche Stimmungsaufhellung. Es gelingt ihr erneut, im Haushalt mitzuhelfen. Sie erreicht allerdings nicht mehr das Ausgangsniveau zum Zeitpunkt vor dem Absetzen des AChEH.

Szenario 2

Bei dem Patienten lag ein Delir bei Alzheimer-scher Erkrankung vor, des Weiteren ein akuter

Harnwegsinfekt in Kombination mit einer Dehydratation. Nach entsprechender Behandlung und vorübergehender Gabe eines Neuroleptikums (wegen der sehr störenden Halluzinationen und der Unruhe) kam der Patient langsam zur Ruhe.

Die berichtete „Unverträglichkeit“ von AChEH war keine absolute Kontraindikation gegen einen erneuten Therapieversuch. Allerdings sprachen bei diesem Patienten die vermehrte Neigung zu Unruhe und deliranter Dekompensation sowie die zusätzliche Bradykardie gegen einen erneuten Therapieversuch. Vielmehr wurde der Patient in ein Pflegeheim mit speziellem Betreuungsprogramm für Demenzkranke entlassen. In dem therapeutischen Milieu sowie unter Sicherstellung optimaler pflegerischer, hausärztlicher und nervenärztlicher Betreuung stabilisierte sich der Patient über längere Zeit erfreulich. Erneute schwerere psychische Begleitsymptome im Rahmen der Demenz traten nicht mehr auf.

23.4 Kommentar

Hanna Kaduszkiewicz, Claus Wächtler

Die Feststellung, dass der wissenschaftliche Nachweis einer Wirksamkeit von AChEH bei Patienten mit Alzheimer-Demenz bisher nicht erbracht wurde, darf nicht mit der Aussage verwechselt werden, dass ihre Unwirksamkeit nachgewiesen worden wäre. Aufgrund der methodischen Mängel der Studien gibt es weder eine Evidenz für ihre Wirksamkeit noch für ihre Unwirksamkeit. Da die Medikamente zugelassen sind, ist auf der individuellen Patientenebene zweierlei Vorgehen möglich: Entweder die „strenge“ Lösung, d.h. der Verzicht auf diese Wirkstoffe, solange der wissenschaftliche Nachweis ihrer Wirksamkeit nicht erbracht ist, oder die „pragmatische“ Lösung, d.h. der Einsatz streng nach Indikation, unter kontinuierlichem Monitoring der Wirkungen und Nebenwirkungen und mit der Option, die Wirkstoffe abzusetzen, wenn sich nach etwa

6 Monaten eine Progredienz der Erkrankung gezeigt hat. Ein Gleichbleiben der Symptomatik wird im Allgemeinen als Erfolg angesehen. Eine unkontrollierte Verschreibung der AChEH sollte der Vergangenheit angehören – wobei über die Modalitäten der Kontrolle, z.B. wer wen kontrolliert, welche Tests mit welchen Cut-Offs eingesetzt werden müssen etc. ein Konsens bisher nicht hergestellt ist. Auf der Systemebene, d.h. bezüglich der Frage, ob diese Medikamente im Leistungskatalog der GKV enthalten sein sollten, ist der Umgang mit der geschilderten Evidenzlage ungleich schwieriger. Aus streng methodischer Sicht sollten sie nicht bezahlt werden, der Kliniker plädiert hingegen unbedingt für eine Bezahlung. Nicht zu vernachlässigen ist in diesem Zusammenhang die Hoffnung auf Hilfe, die Patient und Angehörige aus der Verschreibung eines Medikamentes schöpfen.

Schließlich darf bei allen Diskussionen um die Medikation nicht vergessen werden, dass aus klinischer Sicht gerade Demenzkranke von einem multimethodalen Therapieansatz, also insbesondere auch psychosozialen Maßnahmen, sehr profitieren, auch wenn die Studienlage hierzu **nicht gut** ist [11].

23.5 Zusammenfassung

Hanna Kaduszkiewicz, Claus Wächtler

- ▲ Bei genauer Betrachtung der zu AChEH vorliegenden RCTs zeigen sich sowohl methodische Mängel der Studien als auch Probleme der Anwendbarkeit der Studienergebnisse auf „normale“ Patienten aus dem klinischen Alltag.
- ▲ Aus klinischer Sicht profitiert ein Teil der Patienten von AChEH. Als positive Erfahrungen gelten Stabilisierung des kognitiven Niveaus sowie der Alltagsfertigkeiten über einen Zeitraum von 6–12 Monaten,

Verbesserung nicht-kognitiver Symptome und geringere Belastung der Pflegepersonen. Negative Erfahrungen können neben den typischen Nebenwirkungen z.B. vermehrte Unruhe bei zu Unruhe neigenden Patienten sein.

- ▲ Wer sich für den Einsatz der AChEH entscheidet, sollte dies streng nach Indikation tun, unter kontinuierlichem Monitoring der Wirkungen und Nebenwirkungen und die Wirkstoffe absetzen, wenn sich nach etwa 6 Monaten eine Verschlechterung gezeigt hat. Individuell angepasste psychosoziale Maßnahmen sollten immer erwogen werden.

Literatur

- [1] Müller U et al., Nationale und internationale Demenz-Leitlinien im Vergleich. *Fortschr Neurol Psychiatr.* (2003), 71, 285–295
- [2] Kadoszkiewicz H et al., Cholinesterase inhibitors for patients with Alzheimer's disease: systematic review of randomised clinical trials. *BMJ.* (2005), 331, 321–327
- [3] National Institute for Health and Clinical Excellence (NICE). Donepezil, galantamine, rivastigmine (review) and memantine for the treatment of Alzheimer's disease (appraisal consultation). <http://www.nice.org.uk/page.aspx?o=288826> (01.12.06)
- [4] Rockwood K et al., Attainment of treatment goals by people with Alzheimer's disease receiving galantamine: a randomized controlled trial. *CMAJ.* (2006), 174, 1099–1105
- [5] Winblad B et al., Donepezil in patients with severe Alzheimer's disease: double-blind, parallel-group, placebo-controlled study. *Lancet.* (2006), 367, 1057–1065
- [6] Kadoszkiewicz H, An innovative approach to involve patients in measuring treatment effects in drug trials. *CMAJ.* (2006), 174, 1117–1118
- [7] Olin J, Schneider L, Galantamine for Alzheimer's disease. *Cochrane Database Syst Rev.* (2002), CD001747
- [8] Birks J et al., Rivastigmine for Alzheimer's disease. *Cochrane Database Syst Rev.* (2000), CD001191
- [9] Birks JS, Harvey R, Donepezil for dementia due to Alzheimer's disease. *Cochrane Database Syst Rev.* (2003), CD001190
- [10] Unnebrink K, Windeler J, Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. *Stat Med.* (2001), 20, 3931–3946
- [11] Wächtler C et al., Psychotherapeutische Konzepte bei Demenz. *Psychother Psychiatr.* (2005), 6, 295–303

24 Mammographie-Screening/Prävention – Kritische Bewertung des Screening/Patienteninformation

Ingrid Mühlhauser

24.1 Einleitung/Hintergrund

Die Kontroverse um einen möglicherweise fehlenden Nutzen von Screening auf Brustkrebs bleibt unaufgelöst [1; 2]. Dennoch sollen nun auch in Deutschland alle Frauen im Alter zwischen 50 und 69 Jahren an einem 2-jährlichen Mammographie-Screening teilnehmen. Eine Teilnehmerate von mindestens 70% wird angestrebt. Die Screening-Maßnahme richtet sich an gesunde bzw. beschwerdefreie Frauen. Umso mehr müssen besondere ethische Ansprüche erfüllt werden. Ein „Informed Consent“ spezieller Art ist hier unverzichtbar [3]. International wurden für diesen Informationsprozess Kriterien definiert und ethische Leitlinien publiziert [4–6]. Diese werden zurzeit im deutschen Mammographie-Screening-Programm nicht erfüllt [3].

24.2 Szenarien

Szenario 1

Eine Ihrer langjährigen Patientinnen ist nun 51 Jahre alt geworden. Bisher hat sie noch keine Mammographie-Screening Untersuchung erhalten. Sie möchte wissen, ob sie am Screening teilnehmen soll oder nicht. Sie möchte sicher sein, dass sie keinen Brustkrebs hat. Sie möchte sich nicht später Vorwürfe machen, dass sie nicht alles getan hätte, um das Schicksal einer Brustkrebserkrankung abzuwenden.

Szenario 2

Eine 27 Jahre alte, beschwerdefreie Frau kommt erstmals in Ihre Praxis. Sie möchte wissen, ob sie

nicht auch schon eine Mammographie-Untersuchung durchführen lassen sollte. Die Familienanamnese für Brustkrebs ist unauffällig. Sie hat jedoch eine gleichaltrige Freundin, bei der kürzlich Brustkrebs diagnostiziert und behandelt wurde. Sie ist beunruhigt, da sie so oft liest, dass Brustkrebs jede zehnte Frau trifft. Sie hofft, durch die Vorsorgeuntersuchungen ein solches Schicksal abwenden zu können.

Szenario 3

Eine Ihrer Patientinnen ist nun 76 Jahre alt. Vor fünf Jahren wurde bei einer Screening-Untersuchung ein intraduktales Carcinoma in-situ diagnostiziert, operiert und bestrahlt. Ihre Patientin ist Ihnen zutiefst dankbar, dass der Krebs so früh erkannt wurde, sie fühlt sich geheilt. Sie ist empört, dass Frauen jenseits des 69. Lebensjahres vom Mammographie-Screening-Programm ausgeschlossen sind.

24.3 Kampagnen führen zu Trugschlüssen und Irrtümern

Trugschlüsse und Irrtümer zum möglichen Nutzen und Schaden von Krebsfrüherkennung sind in der Bevölkerung, aber auch in der Ärzteschaft weit verbreitet. So meinen fast 80% der Frauen, dass sie durch Teilnahme am Mammographie-Screening Brustkrebs verhindern oder zumindest das Risiko, daran zu erkranken, vermindern können [7].

Zudem wird das Brustkrebsrisiko überschätzt. Die Darstellung, „jede zehnte Frau trifft es“, ist irreführend. Sie bezieht sich auf die Gesamtlebenszeitspanne und gilt nur für jene Frauen, die das 80. Lebensjahr erreicht

haben. Für Frauen, die gerade das 50., 60. oder 70. Lebensjahr erreicht haben und bisher keine Brustkrebs-Diagnose hatten, gilt, dass bei etwa 1 von 40 in den nächsten 10 Jahren Brustkrebs diagnostiziert wird [8; 9].

Auch der Nutzen des Screenings wird überschätzt. Die Darstellung des möglichen Nutzens als Relativprozent („30% weniger Brustkrebstote“) führt zu einer realitätsfernen Einschätzung dessen, was das Mammographie-Screening leisten kann und was nicht [7].

24.4 Evidenzbasierte Informationen für eine informierte Entscheidung

Wo gibt es brauchbare Informationen für eine informierte Entscheidung? Solche Informationen stehen bisher für Deutschland nur exemplarisch zur Verfügung [10; 11]. Analysen von Internetseiten oder Patientenbrochüren haben immer wieder den Mangel an solchen Informationen deutlich gemacht [5]. Die Fachgesellschaften, Patienten-/Verbraucher-Informationsstellen oder Krankenkassen müssten die Informationen nach EbM-Kriterien – sowohl was den Inhalt als auch was die Art der Präsentation betrifft – verfügbar machen.

24.5 Einschätzung des individuellen Krebsrisikos

Zur Bewertung des möglichen Nutzens und Schadens einer Screening-Intervention sind spezielle Aufbereitungen von Erkrankungs- und Sterbedaten notwendig. Üblicherweise wird jedoch nur das Lebenszeitrisiko angegeben. Es müssten hingegen Angaben über einen für das Individuum relevanten und überschaubaren Zeitraum angeboten werden [9]. Beispielhaft ist dies in Tabelle 25.1 dargestellt. Aktuell gibt es Aufarbeitungen von Daten zu Krebshäufigkeiten nach diesem Schema zu Brustkrebs und Darmkrebs unter Verwendung von Krebsregisterdaten aus Deutschland [8–11; www.gesundheit.uni-hamburg.de].

24.6 Instrumente zur Vorhersage des individuellen Brustkrebsrisikos

Im Internet werden Risikorechner zur Abschätzung des individuellen Risikos für Brustkrebs angeboten (z.B. unter <http://bcra.nci.nih.gov/brc/>). Die Instrumente berücksichtigen neben Alter und familiärer Belastung auch noch andere Risikofaktoren für Brustkrebs, wie z.B. Alter bei Menarche oder Geburt des ersten Kindes. Das bekannteste Instrument ist das so genannte Gail Model. Untersuchungen zur Validierung dieses In-

Tab. 24.1: Wahrscheinlichkeit für Brustkrebs in den nächsten 10 Jahren in Abhängigkeit von der Anzahl von Verwandten 1. Grades mit Brustkrebs; modifiziert nach [12]

Aktuelles Alter der Frau	Anzahl der Verwandten 1. Grades mit Brustkrebs		
	Keine	Eine	Zwei
20 Jahre	4 pro 10 000	1 pro 1000	2 pro 1000
30 Jahre	4 pro 1000	10 pro 1000	20 pro 1000
40 Jahre	14 pro 1000	25 pro 1000	52 pro 1000
50 Jahre	19 pro 1000	32 pro 1000	53 pro 1000
60 Jahre	23 pro 1000	35 pro 1000	56 pro 1000
70 Jahre	25 pro 1000	42 pro 1000	57 pro 1000

struments haben jedoch ergeben, dass eine individuelle Vorhersage für die einzelne Frau nicht möglich ist. Die Vorhersage mit dem Gail Model ist nicht genauer als ein Zufallsergebnis [13]. Zur Abschätzung des individuellen Risikos sind Tabellen, die sich ausschließlich auf das Alter der Frau beziehen, genauso gut oder schlecht wie Instrumente, die mehrere Risikofaktoren berücksichtigen. Auch eine familiäre Häufung ist meistens auf ein zufällig vermehrtes Auftreten von Brustkrebs in der Familie zurückzuführen und nur selten auf eine tatsächlich genetisch bedingte familiäre Belastung.

24.7 Einschätzung des möglichen Nutzens, Schadens bzw. Fehlens von Nutzen

Tabelle 24.2 zeigt exemplarisch solch eine Datendarstellung für das Mammographie-Screening für Frauen zwischen dem 50. und 60. Lebensjahr [2; 14–16]. Neuere Analysen kommen zu ähnlichen Schätzwerten [9; 17; 18].

Tab. 24.2: Beispiel für 50- bis 60-jährige Frauen (Angaben für jeweils 1000 Frauen, Beobachtung über zehn Jahre) (Wahrscheinlichkeitsschätzungen nach dem Krebsregister Hamburg und [2; 8; 9; 14; 15; 18])

	Ohne Mammographie	Mammographie alle 2 Jahre
Verstorben an Brustkrebs	8	6
Verstorben an anderen Todesursachen	72	74
Nicht verstorben	920	920
Diagnose Brustkrebs	25	30
Keine Diagnose Brustkrebs	975	970
Gesamtzahl Screening-Mammographien		5 000
Anzahl Frauen mit mindestens einem falsch positiven Befund		200
Anzahl Frauen mit Biopsie in Folge falsch positiven Befunds		60

Annahmen: Gesamtsterblichkeit 8%, egal ob Mammographie-Früherkennungs-Untersuchungen angeboten werden oder nicht. 30% Verminderung der Brustkrebs-Sterblichkeit durch Mammographie-Früherkennungs-Untersuchungen. Zunahme der Brustkrebs-Diagnosen durch Mammographie-Früherkennungs-Untersuchungen um 20%. Anteil der Frauen mit mindestens einem falsch positiven Mammographie-Befund nach 5 Mammographien 20%. Anteil der Frauen mit Biopsien nach 10 Jahren 6%.

24.8 Auflösung der Szenarien

Szenario 1

Die Wahrscheinlichkeit, dass bei einer beschwerdefreien Frau um die 50, die bisher keine Mammographie hatte, Brustkrebs nachgewiesen werden könnte, liegt bei etwa 0,8% (8 von 1000 Frauen). Die Wahrscheinlichkeit, dass kein Brustkrebs vorliegt, beträgt also etwa 99,2%. Die Wahrscheinlichkeit, dass bei einem negativen Mammographiebefund tatsächlich kein Brustkrebs vorliegt, der sich im Laufe des folgenden Jahres klinisch manifestieren würde, liegt 99,8%. Die Wahrscheinlichkeit, dass kein Brustkrebs vorliegt, ist also auch schon vor Teilnahme an einem Mammographie-Screening sehr hoch. Andererseits besteht auch nach einer Teilnahme am Screening und einem negativen Befund keine 100%ige Sicherheit. Die Patientin sollte die Möglichkeit erhalten, Vor- und Nachteile der Teilnahme am Screening einzuschätzen und eine informierte Entscheidung treffen. Die Broschüre des Nationalen Netzwerks Frauengesundheit wäre eine empfehlenswerte Informationsgrundlage.

Szenario 2

Brustkrebs ist bei Frauen unter 40 Jahre selten (vgl. Tab. 25.1). Durch Screening kann Brustkrebs nicht verhindert und das Risiko, daran zu erkranken, nicht gesenkt werden. Im Gegenteil: Durch Teilnahme am Screening nehmen Brustkrebsdiagnosen zu, es kommt zu so genannten Überdiagnosen und Übertherapien. Die Rate an Überdiagnosen beträgt nach einer Meta-Analyse der randomisiert-kontrollierten Studien zum Mammographie-Screening etwa 30% [2]. Die Patientin sollte die Möglichkeit erhalten, ihr eigenes Risiko besser einschätzen und auch die Nachteile von Screening-Untersuchungen verstehen zu können.

Szenario 3

Nach dem 70. Lebensjahr ist Brustkrebs eine vergleichsweise seltene Todesursache. Wenn vom 70. bis zum 80. Lebensjahr etwa 200 von 1 000 Frauen sterben, sind es nur etwa 6 bis 7 davon, die an Brustkrebs sterben [17]. Kardiovaskuläre und andere Krebserkrankungen führen sehr viel häufiger zum Tod. Das ist ein Grund, warum Screening auf Brustkrebs bei älteren Frauen nicht sinnvoll ist. Zudem ist nicht auszuschließen bzw. sogar wahrscheinlich, dass die Patientin im Szenario eine Brustkrebsdiagnose erhalten hat, die sie ohne Screening nie bekommen hätte. Intraduktale Carcinoma-in situ entwickeln sich nur selten und meist erst nach vielen Jahren zu einem invasiven Brustkrebs.

24.9 Schlussfolgerungen

Das Angebot einer umfassenden, evidenzbasierten und verständlichen Information ist ethisch unverzichtbar und Voraussetzung für jedes Krebsfrüherkennungs-Programm. Die derzeitigen Informationen sind kampagnengeleitet, sie wollen überreden anstatt informieren, motivieren statt aufklären. Die Einladungsschreiben zur Teilnahme am Mammographie-Screening-Programm müssen Informationen bereitstellen, die einen Informed Consent ermöglichen.

Literatur

- [1] Kösters JP, Götzsche PC, Regular self-examination or clinical examination for early detection of breast cancer. *Cochrane Database Syst Rev.* (2003), CD003373
- [2] Olsen O, Götzsche PC, Cochrane review on screening for breast cancer with mammography. *Lancet.* (2001), 358, 1340–1342
- [3] Katzenmeier C, Mammographie-Screening. Rechtsfragen weitgehend ungeklärt. *Dtsch Arztebl.* (2006), 103, A-1054–1058
- [4] General Medical Council (1999) Protecting patients, guiding doctors. Seeking patients' consent: the ethical considerations., London. <http://www.gmc-uk.org>
- [5] Jørgensen KJ, Götzsche PC, Presentation on websites of possible benefits and harms from screening for breast cancer: cross sectional study. *BMJ.* (2004), 328, 148–151
- [6] Steckelberg A et al., Kriterien für evidenzbasierte Patienteninformationen. *Z Arztl Fortbild Qualitätssich.* (2005), 99, 343–351
- [7] Domenighetti G et al., Women's perception of the benefits of mammography screening: population-based survey in four countries. *Int J Epidemiol.* (2003), 32, 816–821
- [8] Mühlhauser I, Höldke B, Information zum Mammographiescreening – vom Trugschluss zur Ent-Täuschung. *Radiologe.* (2002), 42, 299–304
- [9] Kürzl R, Evidenzbasierte Missverständnisse beim Mammakarzinom. *Dtsch Arztebl.* (2004), 101, A-2387–2390
- [10] Nationales Netzwerk Frauen und Gesundheit. Brustkrebsfrüherkennung-Informationen zur Mammographie. Eine Entscheidungshilfe. http://www.gesundheit.uni-hamburg.de/upload/mammografie_1.pdf (01.12.06)
- [11] Mühlhauser I, Höldke B (2000) Mammographie. Brustkrebs-Früherkennungs-Untersuchung. Kirchheim, Mainz. <http://www.mammographie-screening-online.de>
- [12] Collaborative Group on Hormonal Factors in Breast Cancer, Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet.* (2001), 358, 1389–1399

- [13] Rockhill B et al., Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *J Natl Cancer Inst.* (2001), 93, 358–366
- [14] Andersson I et al., Mammographic screening and mortality from breast cancer: the Malmö mammographic screening trial. *BMJ.* (1988), 297, 943–948
- [15] Elmore JG et al., Ten-year risk of false positive screening mammograms and clinical breast examinations. *N Engl J Med.* (1998), 338, 1089–1096
- [16] Nyström L et al., An overview of the Swedish randomised mammography trials: total mortality pattern and the representivity of the study cohorts. *J Med Screen.* (1996), 3, 85–87
- [17] Barratt A et al., Model of outcomes of screening mammography: information to support informed choices. *BMJ.* (2005), 330, 936–938
- [18] Smith-Bindman R et al., Comparison of screening mammography in the United States and the United Kingdom. *JAMA.* (2003), 290, 2129–2137

25 EbM in der Rheumatologie – die interdisziplinäre Leitlinie: Management der frühen Rheumatoiden Arthritis

Monika Lelgemann, Matthias Schneider

25.1 Hintergrund

Klinische Leitlinien können nur dann wirksam werden, wenn die in ihnen enthaltenen Empfehlungen in aktives Handeln umgesetzt werden. Dieses geschieht insbesondere dann, wenn sich die Anwender die Inhalte der Leitlinien zu eigen machen und an die eigenen Bedürfnisse anpassen (adaptieren). Eine wesentliche Voraussetzung für eine erfolgreiche Adaptation und Umsetzung der Empfehlung besteht im Erkennen und Lösen von Implementierungshindernissen, Barrieren, die einer Umsetzung der Leitlinien-Empfehlungen entgegenstehen [1]. Erfolgreiche Beispiele einer solchen Adaptation und Implementierung von Leitlinien existieren im ambulanten Bereich insbesondere für Projekte im Rahmen der Qualitätszirkelarbeit [2].

25.2 Versorgungsproblem

Die seit 12 Jahren bestehende so genannte Kerndokumentation der Deutschen Rheumatologie, gefördert über Programme des Bundesministeriums für Gesundheit und des Bundesministeriums für Bildung und Forschung, hat die in Deutschland bestehenden Defizite in der Versorgung von Patienten mit rheumatoider Arthritis (RA) dezidiert aufgezeigt [3].

Insbesondere besteht ein Verbesserungsbedarf hinsichtlich der frühen Diagnosestellung und der sich ggf. anschließenden Therapieeinleitung. Aggregierte Daten aus Studien belegen, dass ca. 50% aller Patienten mit einer noch undifferenzierten Arthritis und ca. zwei Drittel aller Patienten mit einer frü-

hen rheumatoiden Arthritis im Verlauf von fünf Jahren eine wesentliche Funktionseinschränkung erfahren. Durch den Einsatz von **Disease Modifying Antirheumatic Drugs** (DMARDs), wie z.B. Methotrexat, Sulfalazin oder modernen so genannten Biologicals wie den TNF-Alpha-Antagonisten, mit oder ohne zusätzliche Gabe von Cortison, lässt sich ein Funktionsverlust bei ca. der Hälfte aller Patienten vermeiden und die Chance für eine Krankheitsremission verdreifachen [4]. Diese Angaben gelten dann, wenn die Therapie innerhalb der ersten 6 Monate nach Auftreten der ersten Symptome eingeleitet wird. Auch die Zahl der morphologisch nachgewiesenen Gelenkerstörungen ist mit der Krankheitsdauer bis zur Therapieeinleitung assoziiert.

Das Ziel der multimodalen Therapie der rheumatoiden Arthritis ist es, den Patienten ein Leben ohne wesentliche Behinderungen zu ermöglichen. Wenn es gelingt, durch eine gezielte medikamentöse Therapie eine Gelenkerstörung und damit dauerhaften Schaden zu verhindern, kommt es nicht zu Aktivitätsverlusten und Einschränkungen der Lebensqualität.

Neben der frühen Therapieeinleitung ist hierfür die kontinuierliche Überwachung und ggf. rasche Modifikation der eingeleiteten medikamentösen Therapie entscheidend. Sowohl im Sinne eines Wechsels der primär verordneten DMARDs, im Falle von nicht ausreichendem Ansprechen (Unterdrückung der Krankheitsaktivität) oder auftretenden unerwünschten Wirkungen, aber auch im Sinne eines Absetzens der DMARDs im Falle von nicht mehr nachzuweisender Krankheitsaktivität.

Beide Maßnahmen, die frühe Diagnosestellung und das kontinuierliche Monitoring der Krankheitsaktivität, erfordern die koordinierte Versorgung durch ein multidisziplinäres Team, zu dem primär versorgende Ärzte (Hausärzte, hausärztlich tätige Internisten, Orthopäden), Rheumatologen, Physio- und Ergotherapeuten sowie weitere Fachgruppen gehören.

Gerade Patienten mit einer rheumatoiden Arthritis benötigen von Beginn an eine umfassende Behandlung, welche die individuellen medizinischen, psychosozialen, verhaltenstherapeutischen sowie krankheitsbezogene finanzielle Probleme aufgreift. Die Versorgung sollte in Form einer koordinierten, problemorientierten, multidisziplinären Behandlung erfolgen. Besondere Bedeutung kommt der aktiven Beteiligung der Rheumapatienten an den Therapiestrategien zu.

25.3 Klinisches Szenario – Teil A

In Ihrer hausärztlichen Praxis stellt sich eine Patientin vier Wochen nach ihrem letzten Besuch erneut vor. Damals hatte sie unspezifische Beschwerden in beiden Kniegelenken angegeben, die seit ca. 2 Wochen bestanden. Die Beschwerden waren wechselnder Intensität, es bestand eine leichte Schwellung in beiden Kniegelenken.

Da die Patientin gerade mit ihrem Mann ein altes Haus renovierte, was sie selbst als sehr anstrengend beschrieb, verordneten Sie zunächst ein NSAR bei Bedarf und baten sie, sich bei fortbestehenden Beschwerden erneut vorzustellen.

Die Beschwerden haben sich seit dem letzten Besuch verstärkt, in den Morgenstunden empfindet die Patientin zusätzlich eine eingeschränkte Beweglichkeit der Gelenke. Sie haben den Verdacht, dass es sich um den Beginn einer Erkrankung aus dem rheumatischen Formenkreis handelt und überlegen, welche weiteren diagnostischen Maßnahmen einzuleiten sind.

Sie erinnern sich, dass Sie in Ihrem Qualitätszirkel vor zwei Wochen eine Referentin eingeladen hatten, die ihnen mehrere Internetseiten mit Leitlinien für den hausärztlichen Bereich vorgestellt hat. Unter anderem wurde eine Leitlinie zum Management der frühen rheumatoiden Arthritis vorgestellt, die Sie in dieser Situation zu Rate ziehen möchten.

25.4 Leitlinien als Quelle aufbereiteter Evidenz

Die Leitlinie „Management der frühen rheumatoiden Arthritis“ wurde im Auftrag der Deutschen Gesellschaft für Rheumatologie von einem multidisziplinären Autorengre-


Leitlinienprogramm 				
Klassifikation				
Typ	Charakteristika der Entwicklung	Aufwand	Method. Legitimation	Politische Legitimation
Handlungsempfehlungen von Experten (S1)	Gremium selektiert keine systematische Entwicklung	gering	niedrig	normativ gering
Konsensbasierte LL (S2k)	Gremium repräsent. formalisiertes Konsensverfahren	mittel	niedrig	normativ hoch
Evidenzbasierte LL (S2e)	Gremium selektiert Evidenzbasierung	hoch	hoch	normativ gering
Evidenz- und konsensbasierte LL (S3)	Gremium repräsent. Evidenzbasierung formalisiertes Konsensverfahren	sehr hoch	hoch	normativ hoch

Abb. 25.1: Leitlinien Klassifikationsschema der AWMF (s. auch Kap. 35)

mium erstellt. An der Erstellung waren neben niedergelassenen und stationär tätigen Rheumatologen auch Hausärzte, Internisten, Orthopäden, Physiotherapeuten, Patientenvertreter und Epidemiologen beteiligt.

Es handelt sich nach Klassifikation der Arbeitsgemeinschaft der Medizinisch Wissenschaftlichen Fachgesellschaften (AWMF) um eine S3-Leitlinie [5], also eine Leitlinie, welche die methodischen Anforderungen, die unter anderem im Deutschen Instrument

zur methodischen Leitlinien-Bewertung (DELBI) festgelegt sind [6] (s. Kap. 15, Leitlinien-Bewertung), zumindest weitgehend erfüllt (zur Darstellung der AWMF-Klassifikation s. Abb. 25.1).

Die in der Leitlinie enthaltenen Empfehlungen sind durch Textkästen deutlich gekennzeichnet und mit einem Empfehlungsgrad versehen. Es besteht zudem eine explizite Verknüpfung mit der zugrunde liegenden Literatur (s. Abb.25.2).

Anamnese:		körperliche Untersuchung:	
<ul style="list-style-type: none"> • Morgensteife ≥ 60 Min • Schwellung > 2 Gelenke • Schmerz (Wo?, Was?, Wann?, seit wann?) • allgemeines Krankheitsgefühl 		<ul style="list-style-type: none"> • symmetrisches Verteilungsmuster der betroffenen Gelenke • Schwellung vs. knöcherne Auftreibung/Deformierung • Bewegungseinschränkung (nicht durch andere Ursachen) • extraartikuläre Manifestationen z.B. (Rh.-Knoten) 	
Klinisch richtungsweisender Befund für die Verdachtsdiagnose RA:			
Mehr als 2 betroffene Gelenke seit ≥ 6 Wochen, symmetrisches Verteilungsmuster und Dauer der Morgensteife ≥ 60 Minuten.			
Sind mehr als 2 Gelenkregionen betroffen, bestehen die Entzündungszeichen seit mehr als 6 Wochen und sind sie begleitet von einer morgendlichen Gelenksteife von ≥ 60 Minuten, dann beträgt die Wahrscheinlichkeit des Vorliegens einer persistierenden entzündlichen Gelenkerkrankung etwa 45%. Von diesen 45% werden wiederum 25% erosiv verlaufen.			
Liegt zusätzlich ein positiver Rheumafaktor (IgM-RF ≥ 5 IU) vor, beträgt die Wahrscheinlichkeit einer persistierenden Arthritis 71%, von denen wiederum 50% erosiv verlaufen werden.			
<input checked="" type="checkbox"/>	Bei über 6 Wochen bestehenden Gelenkschwellungen in mehr als 2 Gelenken sollten die Patienten unter dem Verdacht auf eine frühe RA dem Rheumatologen vorgestellt werden.		
A	Die RA sollte von der Diagnosestellung an mit DMARDs behandelt werden, um eine Verzögerung der Krankheitsprogression zu erzielen und damit die Langzeitprognose zu verbessern.	Level 1- → 33	Level 2+ → 21, 22, 25, 33, 34, 35, 182, 188
B	Die Versorgung von Patienten mit rheumatoider Arthritis sollte möglichst von Beginn an in Form einer koordinierten, problemorientierten, multidisziplinären Versorgung erfolgen	Level 1+ → 37, 38	
<input checked="" type="checkbox"/>	Patienteninformation und -schulung sind wesentliche Bestandteile der koordinierten Versorgung von RA-Patienten.		
Die angegebenen Literaturzitate beziehen sich auf die Original-Leitlinie/2. Auflage 2006 http://www.dgrh.de/leitlinien.html			

Abb. 25.2: Kernaussagen der Leitlinien „Management der frühen Rheumatoiden Arthritis“ – Auszug

Die Leitlinie ist über die Leitlinien-Seite der AWMF kostenlos zugänglich. Dort sind neben der Langfassung auch eine Kurzfassung der Leitlinie sowie zusätzliche Materialien zur Information hinterlegt, u.a. eine Patienteninformation, welche die wesentlichen Aussagen der Leitlinie in verständlicher Sprache zusammenfasst und Hinweise auf Patientenorganisationen enthält. Der Prozess der Erstellung ist in einem Methodenreport dargelegt; hier ist ein Gültigkeitsdatum enthalten sowie die Darlegung der Finanzierung der Leitlinienerstellung durch ein durch das Bundesministerium für Bildung und Forschung (BMBF) gefördertes Projekt.

25.5 Klinisches Szenario – Teil B

Nach der Lektüre der Leitlinie (s. Abb.26.2) erhärtet sich Ihr Anfangsverdacht einer Erkrankung aus dem rheumatischen Formenkreis und Sie entscheiden sich im Vertrauen auf die Leitlinie, die Patientin ohne weitere zeitliche Verzögerung z.B. durch eigene Labordiagnostik zum Rheumatologen zu überweisen. Im Aufklärungsgespräch mit der Patientin unterstreichen Sie die Dringlichkeit der fachärztlichen Betreuung. Umso mehr sind Sie erstaunt, am folgenden Tag telefonisch zu erfahren, dass ihre gesetzlich krankenversicherte Patientin erst in 4 Monaten einen Termin beim Rheumatologen erhalten hat und darüber, unter Hinweis auf Ihr gestriges Gespräch, sehr aufgebracht ist.

25.6 Barrieren der Leitlinien-Anwendung

Die Unter- und Fehlversorgung von Patienten mit rheumatoider Arthritis ist in erster Linie auf Schnittstellenprobleme in der Versorgungskette zurückzuführen. Hier steht die Rheumatoide Arthritis stellvertretend für andere, insbesondere chronische Erkrankungen (Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen, 2002).

Besteht das Defizit der Versorgung in erster Linie in einer mangelnden Abstimmung und Koordination aller an der Versorgung beteiligten Disziplinen und Versorgungsebenen, liegt eine Möglichkeit der Verbesserung in der Entwicklung von Modellen der integrierten Versorgung. Hier können Leitlinien, insbesondere wenn sie versorgungsübergreifende Empfehlungen enthalten, als Grundlage der Entwicklung von Konzepten zur Koordination oder z.B. als Grundlage der Entwicklung von die Schnittstellen definierenden Behandlungspfaden dienen [1; 7].

Die oben bereits angesprochene Umsetzung von Leitlinien-Empfehlungen über gemeinsame Adaptation in einem Qualitätszirkel oder Praxisnetz stellt eine weitere Möglichkeit dar. Die damit verbundene Regionalisierung der Leitlinien führt zu einer Anpassung der häufig allgemein gehaltenen Empfehlungen an die konkreten Versorgungsstrukturen und kann dazu beitragen, möglicherweise bestehende Hindernisse der Umsetzung zu beheben [8].

25.7 Diskussion

Neben der Verfügbarkeit der Leitlinien ist ihre effektive Verbreitung eine Grundvoraussetzung dafür, dass Leitlinien zu einer Verbesserung der Versorgung beitragen.

Aus der Literatur ist allerdings bekannt, dass die alleinige (passive) Verbreitung (Diffusion) nicht zu einer Anwendung und Umsetzung der Leitlinien führt, sondern dass seitens der Herausgeber und Autoren aktive Maßnahmen erforderlich sind, um eine Implementierung zu gewährleisten. Unter Implementierung wird dabei der Transfer von Handlungsempfehlungen in individuelles Handeln bzw. Verhalten von Ärzten und anderen Leistungserbringern sowie von Patienten verstanden. Zur Unterstützung dieses Transfers stehen unterschiedliche Maßnahmen zur Verfügung, die seitens der Herausgeber

ber von Leitlinien miteinander kombiniert werden sollten [9]. Diese verschiedenen Maßnahmen können den Kategorien edukativ, finanziell, organisatorisch und regulativ zugeordnet werden. Zielgruppe sollten dabei nicht nur Ärzte, sondern gleichermaßen alle an der Versorgung beteiligten Berufsgruppen und die Patienten sein [10].

Neben aktiven Implementierungsmaßnahmen seitens der „Leitlinien-Ersteller“ ist die Grundbereitschaft von Ärzten, Leitlinien als Informationsquellen zu verwenden, die zweite Voraussetzung dafür, dass Leitlinien wirksam werden können. Nur wenn Leitlinien aktiv als Unterstützung bei der Entscheidungsfindung in spezifischen klinischen Situationen verwendet werden, wenn sie integraler Bestandteil von Aus- und Fortbildung zu den jeweiligen Themen sind, kann ihr Potenzial im Sinne einer Qualitätsverbesserung der Versorgung realisiert werden.

Hierzu gehört auch, Leitlinien als ein Mittel der eigenen Professionalisierung zu verstehen und nicht als Eingriff in die Berufsautonomie, durch den ärztliche Expertise und Erfahrung ersetzt werden sollen.

Grundbedingung für alle geschilderten Formen der Umsetzung und damit auch für die Wirksamkeit von Leitlinien ist dabei deren Validität, wozu selbstverständlich auch die Aktualität gehört. Zweite Voraussetzung ist, dass Leitlinien den tatsächlichen Informationsbedarf decken, was dann besonders gut gelingen wird, wenn die Erfahrungen mit ihrer Umsetzung an die Autoren rückgemeldet werden und somit ein kontinuierlicher Verbesserungszyklus auch für die Leitlinien gewährleistet werden kann.

25.8 Auflösung des klinischen Szenarios

Sie beruhigen zunächst die Patientin und sichern zu, dass Sie den rheumatologischen Kollegen anrufen werden, um für sie einen Termin für eine

kurzfristige Vorstellung zu erhalten. In dem Gespräch mit dem Kollegen stellen Sie gemeinsam fest, dass das Problem struktureller Natur ist, da die Praxis des einzigen niedergelassenen Rheumatologen in der Region völlig überlaufen ist. Sie laden den Kollegen zum nächsten Treffen Ihres hausärztlichen Qualitätszirkels ein und beschließen, auf Grundlage der Leitlinie, ein Konzept zum „Überweisungsmanagement“ von Patienten mit Verdacht auf rheumatoide Arthritis zu erstellen. Dieses soll gewährleisten, dass im Falle eines begründeten Verdachts kurzfristige Termine für die fachärztliche Beurteilung zur Verfügung stehen. Darüber hinaus nehmen Sie Kontakt mit dem regionalen Rheumazentrum auf und verabreden einen Fortbildungstermin für Ihren Qualitätszirkel, auf welchem eine Schulung zum Thema Diagnostik der frühen rheumatoiden Arthritis durchgeführt werden soll. Sie sind sich gemeinsam darüber klar, dass Sie das bestehende strukturelle Problem nicht „unmittelbar“ lösen können, sehen aber anhand der durch die Leitlinie gelenkten Diskussion und der daraufhin verabredeten Maßnahmen eine Möglichkeit, zu einer verbesserten Versorgung der Patienten beizutragen.

Literatur

- [1] Schnabel M et al., Von der Leitlinie zum Behandlungspfad. Entwicklung eines prozessmanagementorientierten Algorithmus zur Akutversorgung polytraumatisierter Patienten. *Chirurg.* (2003), 74, 1156–1166
- [2] Schubert I, PMV-forschungsgruppe. Implementierung hausärztlicher Leitlinien durch Pharmakotherapiezyklen. Erfahrungen aus den Zirkeln der PMV-forschungsgruppe. http://www.pmvforschungsgruppe.de/pdf/03_publicationen/aqua_isi_2003.pdf (01.11.05)
- [3] Zink A, Huscher D, Schneider M, Die Kerndokumentation der Rheumazentren – Bilanz nach 12 Jahren. *Z Rheumatol.* (2006), 65, 144–151
- [4] Schneider M et al. (2005) Management der frühen rheumatoiden Arthritis. Interdisziplinäre Leitlinie. Steinkopff Verlag, Darmstadt

- [5] Kopp I, Encke A, Lorenz W, Leitlinien als Instrument der Qualitätssicherung in der Medizin. Das Leitlinienprogramm der Arbeitsgemeinschaft Wissenschaftlicher Medizinischer Fachgesellschaften (AWMF). Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz. (2002), 45, 223–233
- [6] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI). Fassung 2005/2006. Z Arztl Fortbild Qualitatssich. (2005), 99, 468–519
- [7] Lelgemann M, Ollenschläger G, Evidenzbasierte Leitlinien und Behandlungspfade. Ergänzung oder Widerspruch? Internist (Berl). (2006), 47, 690–698
- [8] Kirchner H, Fiene M, Ollenschläger G, Disseminierung und Implementierung von Leitlinien im Gesundheitswesen: Bestandsaufnahme Juli 2001. Dtsch Med Wochenschr. (2001), 126, 1215–1220
- [9] Grimshaw J, Eccles M, Tetroe J, Implementing clinical guidelines: current evidence and future implications. J Contin Educ Health Prof. (2004), 24 Suppl 1, S31–S37
- [10] Butzlaff M et al., Learning with computerized guidelines in general practice?: A randomized controlled trial. Family Practice. (2004), 21, 183–188

26 Evidenzbasierte Medizin in der Allgemeinarztpraxis

Annette Becker, Norbert Donner-Banzhoff

Evidenzbasierte Medizin und Allgemeinmedizin sind für viele nicht vereinbar. Zu sehr steht die hausärztliche Betreuung für eine Individualmedizin, die neben den biomedizinischen Aspekten einer Behandlung auch den sozialen Kontext eines Patienten und persönliche Eigenschaften berücksichtigt, wie etwa seine Einstellung zu medizinischen Maßnahmen oder ökonomische Faktoren. Die in Studien geprüften Behandlungsbedingungen scheinen willkürlich festgelegt und fernab jeglicher hausärztlicher Realität zu sein. Ist Evidenzbasierte Medizin also für die hausärztliche Praxis untauglich? Anhand eines Fallbeispiels sollen Möglichkeiten der Umsetzung und Grenzen Evidenzbasierter Medizin vor dem Hintergrund typischer allgemeinarztlicher Arbeitsweise dargestellt werden.

26.1 Klinisches Szenario

Während meiner Sprechstunde werde ich telefonisch von Herrn W. um einen Hausbesuch gebeten. Es handelt sich um einen 50-jährigen, mir schon seit langem bekannten und sonst gesunden Patienten, der beruflich stark eingespannt ist und viel reist. Ich erinnere mich, diesen Patienten bereits vor etwa 6 Wochen wegen einer akuten Lumbalgie mit Ibuprofen behandelt zu haben. Am Telefon erscheint mir Herr W. sehr ungeduldig. Man müsse doch endlich herausfinden, warum er diese Schmerzen habe. Außerdem brauche er eine schnell wirkende Spritze – als Unternehmer könne er sich lange Ausfallzeiten einfach nicht leisten.

Ich frage mich, welche Diagnostik ich aufgrund der geschilderten Beschwerden noch

durchführen sollte und welche Behandlung ich dem Patienten nach den Kriterien Evidenzbasierter Medizin empfehlen kann. Bevor ich mich auf den Weg zu Herrn W. mache, möchte ich mir einen Überblick über diagnostische und therapeutische Verfahren verschaffen.

26.2 Die Evidenz suchen

Die Fragestellung für den vorliegenden Fall muss allgemein formuliert werden und könnte folgendermaßen lauten: „Welche Diagnostik sollte man diesem 50-jährigen Patienten mit Kreuzschmerzen zukommen lassen und welche Therapien können effektiv die Beschwerden des Patienten lindern bzw. die Rezidivrate senken?“

Meist ist die zur Verfügung stehende Zeit im Praxisalltag zu knapp, um eine aufwendige Literatursuche durchzuführen und Therapiestudien für alle Aspekte des Beratungsanlasses zu sichten. Der Hausarzt muss unter den ihm zur Verfügung stehenden Quellen eine aussuchen, mit der er die Frage beantworten kann. Dabei stehen ihm nicht nur die klassischen Quellen zur Evidenzbasierten Medizin (MEDLINE, Cochrane, Clinical Evidence und Leitlinien) zur Verfügung, sondern auch Arzneimitteldatenbanken wie die Gelbe Liste, Fachinformationen der Hersteller, die Rote Liste und andere. Jede dieser Quellen hat ihre eigene Indikation. So kann es sinnvoll sein, bei fokussierten Fragen, die konkrete Verfahren im Vergleich zu Alternativtherapien betreffen (z.B.: Massage vs. Bewegungstherapie bei Rückenschmerzen) auf die MEDLINE zuzugreifen. Für Therapiefragen im Allgemeinen bie-

ten sich Clinical Evidence, die Cochrane Library oder bei medikamentösen Fragestellungen neutrale Arzneimittelinformationen wie z.B. das Arzneimitteltelegramm an. Handelspräparate, Dosierungen oder Wechselwirkungen lassen sich gut in der Roten Liste, der Gelben Liste oder den Fachinformationen der Hersteller recherchieren.

Im vorliegenden Fall ist die Fragestellung sehr wenig differenziert. Eine Suche in MEDLINE ist hierfür nicht geeignet. Dort finden sich zum Begriff „back pain“ über 24 000 Einträge. Selbst weiteres Eingrenzen der Suche durch zusätzliche Suchbegriffe wie „low back pain“, „therapy“ oder die Beschränkung auf systematische Reviews reduziert die Trefferzahl nicht so weit, dass man sich zügig über gängige Therapiemethoden informieren könnte. Es kommt schnell zu einem Gefühl der Überforderung angesichts der Flut an Informationen in englischer Sprache. Stattdessen bietet sich der Rückgriff auf evidenzbasierte Leitlinien an, wo man schnell die Diagnostik und Therapie verschiedener Beratungsanlässe überblicken kann. Auf Leitlinien zum Thema Rückenschmerz trifft man auf der Homepage der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF, <http://www.awmf-leitlinien.de>); darunter findet sich auch eine Leitlinie Kreuzschmerzen der Deutschen Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM), die in Teilen online verfügbar ist (http://www.degam.de/leitlinien/3_kreuzschmerzen.html).

26.3 Die Evidenz bewerten

Die erste Durchsicht der Leitlinie dient dazu, sich einen Überblick über die Methodik der Leitlinienerstellung zu verschaffen, um die Validität der Empfehlungen einschätzen zu können.

Die gewählte Leitlinie hat ein interdisziplinäres Team aus Wissenschaftlern und prak-

tisch tätigen Ärzten erstellt. Sie gibt einen breiten Überblick über verschiedene Aspekte der Diagnostik und Behandlung von erwachsenen Patienten mit akuten und chronischen Kreuzschmerzen. Aktivitätseinschränkende Beschwerden von weniger als drei Monaten Dauer werden als akut definiert, Schmerzen seit drei Monaten und mehr als chronisch. Invasive Verfahren werden ebenso berücksichtigt wie pharmakotherapeutische, physikalische, gesprächs- und verhaltenstherapeutische Maßnahmen. Die Evidenzlevel umfassen „höchste“ Evidenz (A, z.B. Metaanalysen und randomisierte kontrollierte Studien bei therapeutischen Fragen), „mäßige“ Evidenz (B, z.B. Kohortenstudien, Fall-Kontroll-Studien, Querschnittstudien oder Fallserien für Therapiestudien) und „schwache“ Evidenz (C, Expertenmeinung, Grundlagenforschung). Es werden Publikationen berücksichtigt, die die Linderung von Schmerzen, funktionelle Einschränkungen oder die Rückkehr an den Arbeitsplatz untersuchen (Outcome). Die Leitlinie ist 2003 veröffentlicht worden und soll 2006 revidiert werden.

Ich bin mit der Validität der Leitlinie zufrieden, sodass ich mich den inhaltlichen Aussagen zuwende: Zur Diagnostik von Kreuzschmerzen empfiehlt die Leitlinie eine „diagnostische Triage“. Anhand von Anamnese und klinischer Untersuchung wird unterschieden zwischen

- ▲ unkomplizierten Kreuzschmerzen (bewegungsabhängig, eventuell dermatomübergreifend mit Ausstrahlung bis oberhalb des Knies, guter Allgemeinzustand),
- ▲ radikulären Kreuzschmerzen (einseitige Schmerzen im Bein, ausstrahlend bis unterhalb des Knies, gegebenenfalls mit Taubheitsgefühl und Parästhesien, positiver Lasègue-Test, Reflexauffälligkeiten) und
- ▲ komplizierten Kreuzschmerzen bei Warnhinweisen. Als Warnhinweise für abwendbar gefährliche Verläufe („Red Flags“) gelten z.B. ein Alter unter 20 und

über 55 Jahren, Frakturen, Tumore oder Entzündungen. Liegen diese vor, ist eine weiterführende Diagnostik zu erwägen. Andernfalls soll auf bildgebende Verfahren verzichtet werden.

Therapeutisch empfehlen die Autoren der Leitlinie als Basistherapie eine eingehende Beratung über die Harmlosigkeit der Beschwerden, Vermeidung von Bettruhe und die Notwendigkeit körperlicher Aktivität sowie eine Analgesie, vorzugsweise mit nicht-steroidalen Antirheumatika oder Paracetamol. Bei akuten Beschwerden kann eine chirotherapeutische Behandlung erwogen werden; rezidivierende und persistierende Beschwerden mit Arbeitsunfähigkeitszeiten von weniger als 3 Monaten profitieren nach erneuter Evaluation der Beschwerden von einer Intensivierung der bisherigen Therapie, Physiotherapie, Rückenschulen oder gegebenenfalls Massagen. Liegen Risikofaktoren für chronische Verläufe vor, wird Psychotherapie empfohlen. Chronische Rückenschmerzen und Beschwerden bei Arbeitsunfähigkeitszeiten von drei Monaten und mehr sollten multiprofessionell in Zusammenarbeit von z.B. Physiotherapeuten, Schmerztherapeuten, Orthopäden und Psychotherapeuten in entsprechenden Zentren betreut werden. Alle anderen in der Leitlinie diskutierten Therapien zeigen entweder keine oder nur mangelhafte Evidenz, weshalb sie nur zweite Wahl sind und möglichst nur dann eingesetzt werden sollten, wenn sie den aktivierenden Therapieansatz unterstützen.

26.4 Klinisches Szenario

Herr W. öffnet mir die Tür. Er bewegt sich vorsichtig in leicht vorgebeugter Haltung und gibt an, seit etwa vier Tagen an starken Kreuzschmerzen zu leiden. Im Gespräch erfahre ich, dass Herr W. seit der letzten Behandlung beschwerdefrei gewesen sei. Vor einigen Tagen habe es aber wie aus

heiterem Himmel wieder angefangen. Er sei von einer langen Dienstreise im PKW wiedergekommen und habe bereits gegen Ende der Fahrt starke Schmerzen im Bereich der Lendenwirbelsäule entwickelt. Er meide zurzeit Bewegung, da Positionsänderungen, wie z.B. aus dem Sitzen aufzustehen, seine Schmerzen verschlimmerten und teils in den rechten Oberschenkel ausstrahlten.

Meine Untersuchung ergibt neben einer Schonkyphose keine Fehlhaltung und keinen Hinweis auf einen Beckenschiefstand. Herr W. gibt Schmerzen im Bereich L4-5/S1 an, die bei aktiver Anteversion der Wirbelsäule schlimmer werden (Finger-Boden-Abstand ca. 40 cm). Ich taste einen Muskelhartspann beidseits paravertebral im Bereich L3-L5. Das passive Anheben jeweils eines der gestreckten Beine verursacht ihm lumbal Schmerzen ohne Ausstrahlung. Es bestehen keine Parästhesien, Sensibilitäts- oder Reflexausfälle der unteren Extremitäten.

Ich stelle die Diagnose rezidivierender unkomplizierter Rückenschmerzen und fühle mich darin bestärkt, dass vorerst keine weitere Diagnostik durchgeführt werden muss. Herrn W. erläutere ich die Befunde, erkläre die Gutartigkeit des Krankheitsbildes und die hohe Wahrscheinlichkeit, dass seine Beschwerden innerhalb der nächsten Tage vollständig verschwinden werden. Ich erkläre ihm, dass eine Spritze zur Behandlung seiner Beschwerden nicht sinnvoll sei. Da er bislang unregelmäßig Paracetamol genommen hat, empfehle ich ihm stattdessen Diclofenac 75 mg zweimal täglich für 2 Tage. Im Gespräch erfahre ich, dass Herr W. bereits aufmerksam darauf geworden ist, dass Rückenschmerzen mit Bewegungsmangel zusammenhängen, konkrete Umsetzungspläne für sich selbst hat er allerdings nicht und ist heute auch nicht gewillt, sich damit auseinanderzusetzen. Gemeinsam entwickeln Herr W. und ich die Strategie, die Beschwerden vorerst analgetisch zu behandeln. Ich rate ihm, seine Arbeit morgen oder am Folgetag wieder aufzunehmen, eine erneute Dienstreise in diesen Tagen aber zu verschieben. Außerdem biete ich ihm an, in naher Zukunft Strategien zur Vermeidung weiterer Rezidive mit ihm zu besprechen.

26.5 Diskussion

Rückenschmerzen sind ein typisches Problem in hausärztlichen Praxen, für das ein Überangebot an diagnostischen und therapeutischen Möglichkeiten besteht. Die Therapie variiert von minimalistischen Behandlungsansätzen bis zum Einsatz hochtechnisierter Geräte. Bei hoher Spontanheilungsrate kommt es häufig zu Rezidiven, und bis zu 10% der Patienten entwickeln einen chronischen Verlauf. Die Prognose der Beschwerden wird entscheidend durch das ärztliche Vorgehen beeinflusst, denn übertriebene Diagnostik und passive Therapien fördern die Chronifizierung der Beschwerden, da sie das Krankheitserleben der Patienten verstärken und einer frühzeitigen Aktivierung entgegenstehen. Der Hausarzt steht im Spannungsfeld zwischen Heilungswunsch, den Erwartungen des Patienten, den Auswirkungen übertriebener Diagnostik und Therapie auf den Heilungsprozess und externen Zwängen wie ökonomischen Begrenzungen oder fehlenden Behandlungsplätzen.

Das vorliegende Problem kann relativ zügig recherchiert und anhand der Leitlinie Kreuzschmerzen der DEGAM bearbeitet werden. Jedoch: Nicht immer lassen sich allgemeinärztliche Probleme so leicht wie im vorliegenden Fall lösen. Bislang gibt es nur wenige allgemeinmedizinische Leitlinien, die in ihrem Aufbau und ihren Empfehlungen den hausärztlichen Gegebenheiten entsprechen. Dagegen sind die Inhalte der Leitlinien anderer Fachgesellschaften oder anderer EbM-Quellen oft nicht auf das hausärztliche Umfeld und die allgemeinärztliche Arbeitsweise übertragbar.

Ein Grund hierfür ist, dass sich hausärztliche Patienten von denen in spezialisierten Einrichtungen unterscheiden. Einerseits betreuen Hausärzte viele Patienten in höherem Lebensalter, chronisch erkrankte, multimorbide oder sozial schwache Patienten, die – wenn sie nicht im Design einer Studie von

vornherein ausgeschlossen werden – die „Randgruppen der Studienpopulationen“ darstellen. Andererseits arbeiten Hausärzte im Bereich der Niedrigprävalenz: Sie werden als häufig erste Anlaufstelle des Versorgungssystems mit vielen, aus medizinischer Sicht „banalen“ Problemen konfrontiert. Die Wahrscheinlichkeit für eine zugrunde liegende schwerwiegende Erkrankung ist gering, sodass bei Spezialisten angewandte Strategien zur Differenzialdiagnose oder anscheinend wirksame Therapien in der Hausarztpraxis nicht zu gleichen Ergebnissen führen. Ein aktuelles Beispiel hierfür ist die Diskussion um das Ausmaß der therapeutischen Blutdrucksenkung: Es bestehen Zweifel, ob die in großen Studien (an jüngeren Patienten) beobachteten positiven Effekte einer straffen Blutdrucksenkung (auf 130/80 mmHg) auch auf primärmedizinische Patienten zu übertragen sind oder ob sie nicht stattdessen die Gesamtmortalität der Patienten erhöhen.

Trotz der meist unkomplizierten Beschwerden müssen die wenigen abwendbar gefährlichen Verläufe sicher erkannt werden. Tests, die in einer spezialisierten Ambulanz einen guten prädiktiven Wert haben, verlieren aufgrund der Niedrigprävalenz in Hausarztpraxen ihre Vorhersagekraft. Ein praktikables Konzept ist die Arbeit mit Warnhinweisen oder Red Flags, wie im Beispiel des Patienten mit Rückenschmerzen. Während die Prävalenz abwendbar gefährlicher Verläufe bei Rückenschmerzen mit 1%–3% sehr gering ist und aus diesem Grund keine weitere Diagnostik indiziert ist, steigt die Wahrscheinlichkeit für das Vorliegen einer zugrunde liegenden Pathologie mit dem Zusammenkommen verschiedener Risikofaktoren, eben dieser Warnhinweise. Liegen Warnhinweise vor, steigt die Vortest-Wahrscheinlichkeit und eine vorher noch abzulehnende bildgebende Diagnostik wird sinnvoll.

Diese symptomorientierte Arbeitsweise ist typisch für die Hausarztpraxis. Statt klarer

Diagnosen stellt sich dem Hausarzt häufig ein undifferenziertes Bild von Symptomen, die zunächst nicht richtungsweisend sind und anfangs mit abwartendem Offenhalten behandelt werden. Eine besondere Bedeutung haben in diesem Zusammenhang symptomevaluierende Studien, die Aufschluss über die Häufigkeitsverteilungen von Symptomen, die Wahrscheinlichkeit für zugrunde liegende Erkrankungen (Differentialdiagnose) und die Prognose der Patienten mit diesen Symptomen untersuchen. Die Antworten auf diese Fragen sind abhängig von der Rekrutierungsmethode (z.B. wird nach Beratungsanlass oder dem Vorhandensein von Symptomen unabhängig vom Konsultationsgrund rekrutiert), dem Studien-Setting (z.B. Krankenhausambulanz oder Allgemeinarztpraxis) bzw. der Bezugsgröße, dem Nenner („Beziehen sich die relativen Daten auf die Gesamtbevölkerung oder alle hausärztlichen Patienten?“). Ätiologische oder prognostische Studien sollten Vergleichsgruppen von Patienten ohne das Symptom untersuchen und zwar mit ausreichend langem Beobachtungszeitraum (Follow-up). Häufig ist es notwendig, mehrere Zentren oder Praxen in Studien einzubeziehen, um auf eine ausreichend hohe Fallzahl zu kommen. Tabelle 26.1 fasst die Qualitätskriterien

zur Beurteilung der Validität symptomevaluierender Studien noch einmal zusammen.

Was ist aber zu tun, wenn die externe Evidenz nicht ausreicht? Im Beispiel Rückenschmerz sind viele Therapieoptionen nicht ausreichend evaluiert. Stattdessen gibt es starke Evidenz, dass weitere bildgebende Diagnostik und die Injektion von Analgetika, die der Patient wünscht, nicht angebracht sind und dem Verlauf der Erkrankung sogar schaden können. Der Allgemeinarzt ist gefordert, dem Patienten einerseits etwas abzuschlagen, andererseits ihn zur Mitarbeit bzw. zu mehr Aktivität zu ermuntern. Die Situation ist schwierig und erfordert ein hohes Maß an kommunikativen Fähigkeiten. An dieser Stelle kann die externe Evidenz die Basis für den Handlungsspielraum geben. Wie die Umsetzung aber konkret erfolgt, bleibt dem behandelnden Arzt überlassen. Die lange Arzt-Patient-Beziehung mit Kenntnis der persönlichen Präferenzen, des Umfeldes und der Bedürfnisse seines Patienten sowie das Wissen um logistische Fragen, Kooperationsmöglichkeiten oder Ähnliches (interne Evidenz) können zu einem (im Vergleich zu den Leitlinienautoren) veränderten Wertesystem führen und die Entscheidung beeinflussen. Es obliegt dem Arzt, die Empfehlungen der Leitlinie auf individuelle Gegebenheiten zu

Tab. 26.1: Qualitätskriterien zur Beurteilung symptomevaluierender Studien

Prospektive Rekrutierung der Studienteilnehmer
Einschluss aller konsekutiv vorstellig gewordenen Patienten
Klare Definition der Einschlusskriterien
Multizentrische Stichprobe mit hoher externer Validität
Klare Definition der Bezugspopulation
Vergleichsgruppe von Patienten ohne das untersuchte Symptom
Transparente, einheitliche und standardisierte Untersuchungen
Adäquate diagnostische Kategorien
Ausreichend langes und komplettes Follow-up
Ausreichende Fallzahl für präzise Schätzungen der wichtigsten Variablen und Präsentation von Konfidenzintervallen

übertragen. So kann es in Einzelfällen durchaus sinnvoll sein, zunächst auf Therapien unzureichender Evidenz zurückzugreifen, um sich im Entscheidungsprozess dem Patienten anzunähern und Raum für eine schrittweise Hinführung zur evidenzbasierten Therapie zu schaffen.

27 Individualisierte Risikoproggnose für die Herz-Kreislauf-Prävention: ein Werkstattbericht

Norbert Donner-Banzhoff, Heidemarie Keller, Eva-Maria Sadowski, Tanja Krones, Erika Baum, Andreas C. Sönnichsen, Uwe Popert

Diskussionen um die Evidenz beziehen sich meist auf Therapie-Studien – das gilt für EbM-Kurse wie auch für gesundheitspolitische und -ökonomische Kontroversen. Den einzelnen Patienten und seine Familie bewegt in mindestens gleichem Maße die Frage nach der Prognose. Dies mag sich auf manifeste Erkrankungen beziehen (z.B. Lebensdauer und -qualität bei malignen Erkrankungen, Arbeitsfähigkeit bei Rückenschmerzen), aber auch auf asymptomatische Risikozustände (z.B. die Hypertonie).

Für die Gewichtung kardiovaskulärer Risikofaktoren sind in den letzten 15 Jahren Instrumente zur individualisierten Prognose entwickelt worden. Diese spiegeln ein verändertes Behandlungsparadigma wieder: individuelle Risikofaktoren (z.B. der Blutdruck) werden nur noch bei extremen Abweichungen „korrigiert“, ansonsten steht das kardiovaskuläre Gesamtrisiko im Vordergrund. Darin stimmen sämtliche aktuellen Behandlungsleitlinien überein [1].

Wir demonstrieren hier den Entwicklungsprozess eines Beratungsinstruments für die hausärztliche Praxis, das diese Entwicklungen aufnimmt und in eine umfassend implementierbare Form bringt. Damit lässt sich eindrücklich zeigen, welche Hindernisse, Zwänge, aber auch Chancen sich bei der Anwendung von Studienevidenz im Alltag der Versorgung ergeben.

27.1 Die Beratungsstrategie „ARRIBA-Herz“

Der einzelne Patient stellt sich die Frage: Wie hoch ist mein Risiko, einen Herzinfarkt oder einen Schlaganfall zu erleiden? Und welche präventiven Maßnahmen sind in meinem Fall angemessen und sinnvoll? In seiner Antwort muss der Arzt mehrere Risikofaktoren gleichzeitig berücksichtigen, was in Bezug auf Berechnung und Darstellung ausgesprochen schwierig ist. Außerdem sollen Risikoprognose und die Effekte möglicher präventiver Maßnahmen so dargestellt werden, dass der Patient sich eine eigene Meinung bilden und mitentscheiden kann.

Um diesen Anforderungen gerecht zu werden, haben wir seit 2001 ARRIBA-Herz entwickelt und seitdem fortwährend angepasst und evaluiert. Konkret liegt eine Broschüre vor, die eine entsprechende Beratung (s. Tab. 27.1) ausführlich erläutert; darin werden epidemiologische, medizinische und ethische Hintergründe diskutiert sowie konkrete Hilfen zur Gesprächsführung gegeben. Als Umsetzungshilfe werden ein Risiko-Kalkulationsbogen für Frauen und Männer zur Verfügung gestellt, der auch einen Vergleich mit der nach Alter und Geschlecht definierten Bevölkerung bietet. Eine visuelle Darstellung („Smilies“) des 10-Jahres-Risikos kann auf einem weiteren Bogen dokumentiert werden; hier sind auch präventive Möglichkeiten aufgezeigt, entsprechende Behandlungspläne können dokumentiert werden. Schließlich haben wir für die weitere Verbreitung ein Moderatorenmanual und eine Powerpoint-Präsentation erstellt.

Tab. 27.1: ARRIBA-Schritte

Aufgabe gemeinsam definieren: individuelles Gesamtrisiko bestimmen, Patient in die Entscheidung einbeziehen
Risiko subjektiv: Befürchtungen, Erwartungen, Wünsche, Fragen des Patienten
Risiko objektiv: messen und Punkte addieren, in Wahrscheinlichkeit umrechnen und mit Bevölkerung vergleichen; Wahrscheinlichkeit optisch darstellen
Information über Möglichkeiten zur Vorbeugung (Prävention): optische Darstellung absoluter Effekte
Beiderseitige Bewertung: Alternativen, Vor- und Nachteile
Absprache über das weitere Vorgehen: Prioritäten setzen – Verhaltensänderungen, Medikamente, Kontrollbesuche

Seit fünf Jahren werden die Rückmeldungen von Hausärzten systematisch ausgewertet, die an Fortbildungen zu ARRIBA-Herz teilgenommen haben. Mit diesen Hinweisen ist das Instrument fortlaufend verbessert und den Bedürfnissen der hausärztlichen Praxis angepasst worden. Dieser Entwicklungsprozess kulminiert in einer Phase-III-Studie, die eine Überlegenheit der ARRIBA-Beratung im Vergleich zum herkömmlichen Vorgehen zeigen konnte [2].

Neben den kritischen Rückmeldungen der Zielgruppe sind neue Erkenntnisse aus epidemiologischen und Präventionsstudien in die Weiterentwicklung eingeflossen. Letztere verlangen allerdings umfangreiche Überlegungen und Entscheidungen, da sich die rohen Ergebnisse von Studien meist nicht unmittelbar in die Beratungssituation mit dem einzelnen Patienten übertragen lassen.

27.2 Entscheidungen

Die Umsetzung epidemiologischer Erkenntnisse in die Praxis erzwingt immer einen Kompromiss von Praktikabilität, Validität und Präzision. Dabei ist das erste Erfordernis grundsätzlich zu erfüllen, da sonst das Instrument überhaupt keine Wirkung entfalten wird. Zur Praktikabilität gehören Umsetzbarkeit im Alltag, Zeitökonomie sowie Verständlichkeit und Plausibilität. Validität

bedeutet, dass die Ergebnisse bestmöglicher aktueller Studien berücksichtigt werden; hier sind sowohl die interne (z.B.: Sind präventive Effekte durch RCTs abgesichert?) wie auch die externe Validität (z.B.: Für welche ethnischen Gruppierungen ist ein Prognoseinstrument gültig?) zu bedenken. Natürlich wird angestrebt, eine individuelle Risikoprognose mit größtmöglicher Präzision zu erstellen, wobei individuelle Gegebenheiten (Risikofaktoren) berücksichtigt werden.

Vor diesen sich vielfach widersprechenden Anforderungen gibt es in der Regel keine ideale Lösung. Lösungen sind vielmehr das Ergebnis sorgfältiger Abwägungen und praxisbasierter Erprobung.

27.2.1 Auswahl einer prädiktiven Formel

Inzwischen gibt es drei multivariate Formeln, die jeweils an großen, über mehrere Jahre verfolgten Kollektiven erstellt wurden. Diese machen es möglich, aus einzelnen Risikofaktoren eine individuelle Risikoprognose zu berechnen; alle drei sind jeweils an anderen Stichproben validiert worden. Es handelt sich um die Framingham- [3–5], PROCAM- [6] und SCORE- [7] Instrumente.

Wir haben uns entschieden, bei ARRIBA die Framingham-Formel zu benutzen. Eine repräsentative Bevölkerungsstichprobe, mehrfache Validierung an anderen Kollektiven, ein

präziser Schätzer für Frauen, für Patienten mit antihypertensiver Therapie und solche bis zu einem Alter von 79 Jahren waren die Gründe. Eine kürzlich publizierte Version [5] liegt den Empfehlungen des National Cholesterol Education Program (NCEP) ATP-III-Panels zu Grunde [8]. Obwohl das PROCAM-Instrument auf einer deutschen Stichprobe beruht, ist seine externe Validität problematisch (in Behörden und größeren Unternehmen beschäftigte Männer von 40–65 Jahren; geringer Validierungsgrad in Bezug auf andere Kollektive). Die Notwendigkeit der Triglyzerid-Bestimmung bringt zusätzlichen Aufwand mit sich und ist möglicherweise in den spezifischen Interessen der Autoren begründet. Die SCORE-Prognosen beziehen sich ausschließlich auf die Mortalität; die berücksichtigten Risikofaktoren haben ein Element von Willkür, da sie von der Daten-Verfügbarkeit der einzelnen Studien abhängen. Zudem liegt für das SCORE-Instrument kein Punktesystem vor, das bei ARRIBA-Herz die Voraussetzung für die visuelle Darstellung ist (s. u.).

Ein Nachteil der Framingham-Formel ist allerdings, dass sie die Erkrankungswahrscheinlichkeit für Europäer überschätzt [9; 10] (s.u.). Dazu kommt das säkulare Absinken der kardialen Morbidität und Mortalität; letzteres ist ein Argument dafür, möglichst aktuell entwickelte Prognose-Instrumente einzusetzen.

27.2.2 Zielkriterium

Bei der Erprobung unserer ersten Version, die sich lediglich auf die kardiale Morbidität und Mortalität bezog, ergab sich, dass eine Prognose in Bezug auf den Schlaganfall von Patienten und Ärzten als ebenso wichtig erachtet wurde. Zwar liegt eine von der Framingham-Kohorte abgeleitete Risikoformel für Schlaganfälle vor [11]; der Einsatz von zwei Berechnungen (kardiale und zerebrovaskuläre Ereignisse getrennt) erschien uns aller-

dings als nicht praktikabel. Wir haben deshalb einen Aufschlag auf das kardiale Risiko berechnet, der aus dem Verhältnis von kardialen zu zerebrovaskulären Erkrankungen in den europäischen MONICA-Studien abgeleitet ist [12]. Die Summe aus dem mithilfe der Framingham-Formel berechneten kardialen Risiko und dem Aufschlag für zerebrovaskuläre Ereignisse spiegelt somit das kardiovaskuläre Gesamtrisiko wider.

27.2.3 Subjektivität

Begriffe wie „Risiko“ oder „Wahrscheinlichkeit“, auch „Herzinfarkt“ oder „Schlaganfall“ führen bei Patienten zu sehr unterschiedlichen Vorstellungen, die von dem epidemiologisch bzw. medizinisch Gemeintem stark abweichen können. Nicht nur der Informationsstand (Ätiologie, Prognose, Behandlung), sondern auch assoziierte Befürchtungen und Ängste spielen eine Rolle, die wiederum auf subjektive Erfahrungen zurückgehen. Das hausärztliche Beratungsgespräch bietet die Möglichkeit, das Verständnis fortlaufend zu kontrollieren und die Informationsvermittlung individuell anzupassen. Eine Kenntnis der familiären Situation ist hier besonders wertvoll; so mag ein Patient unter dem Eindruck eines gravierenden Schlaganfalls in der Familie stehen und deshalb nach präventiven Möglichkeiten greifen, die rational gesehen nur unzureichende Effekte bieten. Hier kann der Hausarzt die Befürchtung des Patienten direkt ansprechen und beispielsweise verhindern, dass momentane Eindrücke ein zu großes Gewicht auf eine Entscheidung bekommen. Die Erörterung der subjektiven Aspekte des kardiovaskulären Risikos ist deshalb ein obligatorischer Schritt in der ARRIBA-Herz-Beratung. Dies ist ein wichtiger Unterschied gegenüber einem Internet-basierten Risikorechner, dessen Resultat der Patient für sich allein interpretieren und gewichten muss.

27.2.4 Zeitdauer

In der ersten Publikation des Framingham-Instruments [3] wurden Angaben für einen 5- und einen 10-jährigen Bezugszeitraum gemacht. Da die einschlägigen Interventionsstudien nicht über fünf Jahre hinausgingen, haben wir in der ARRIBA-Herz-Beratung zunächst nur diesen Zeithorizont berücksichtigt. Die neueren Versionen des Instruments [5] – wie übrigens auch SCORE – beziehen sich allerdings auf einen Zeitraum von zehn Jahren. Ein Grund ist sicher die schwierige Darstellung „mikroskopischer“ Risiken v.a. bei jüngeren Menschen; andererseits ist ein 10-Jahres-Risiko zwangsläufig „höher“ und kann deshalb vermehrt Ängste verursachen. Hier werden die subtilen Manipulations-Gefahren deutlich; mehrfach ist nachgewiesen worden, dass verschiedene Arten der Risikokommunikation („Framing“) jeweils zu unterschiedlichen Entscheidungen führen können [13].

27.2.5 Regionale Anpassung

Da wir auch bei der neuesten Framingham-Formel von einer Überschätzung des Risikos für europäische Stichproben ausgehen müssen, adjustierten wir mit einem Kalibrierungs-Faktor (Multiplikation mit 0.68) auf der Basis der British Regional Heart Study [9]. In derselben Größenordnung bewegt sich die Überschätzung des Risikos an der deutschen PROCAM-Kohorte durch die Framingham-Formel [10]. Allerdings beruhen diese Berechnungen auf älteren Framingham-Algorithmen, neuere Berechnungen liegen nicht vor.

27.2.6 Bevölkerungsvergleich

Für den einzelnen Patienten ist die Interpretation einer Risikoprognoze schwierig. Eine Hilfe kann dabei der Vergleich mit einer geschlechts- und alters-identischen Bevölkerungs-Stichprobe sein. Das Robert-Koch-Institut war freundlicherweise bereit, aus einer repräsentativen deutschen Bevölkerungsstichprobe (Bundesgesundheits-Survey 1998) die Perzentilenwerte der aktuellen Framingham-Formel zu berechnen, sodass sich in der ARRIBA-Beratung jeder Patient entsprechend einordnen kann.

27.2.7 Diabetes

Die Hypothese, dass ein Diabetes mellitus Typ II ein ebenso hohes kardiovaskuläres Risiko bedingt wie eine manifeste KHK [14], lässt sich nach zahlreichen weiteren Publikationen zu dieser Fragestellung nicht mehr aufrecht erhalten [15–17]. Unglücklicherweise gehen die Autoren der aktuellen Framingham-Formel (Punktesystem) noch von dieser Annahme aus und haben deshalb den Diabetes nicht berücksichtigt.

Um den Diabetes und auch den Grad seiner Kontrolle einzubringen, beinhaltet ARRIBA deshalb einen Faktor, der auf einer Meta-Analyse von 10 Kohorten mit Diabetes mellitus Typ 2 beruht [18]. Dabei ergab sich ein Relatives Risiko von 1.18 für kardiovaskuläre Ereignisse bei einer Steigerung des HbA1c um einen Punkt. Bei einem Referenzwert von HbA1c=6.0 ergibt sich für die Mittelwerte der von uns zur Vereinfachung gewählten Klassen (6–7.9; 8–10) jeweils eine Erhöhung des Risikos für relevante Ereignisse um 20% bzw. 60%.

Eine separate Formel für Diabetiker (UKPDS-Risk-Engine [19]) hätte wegen des höheren Aufwands die Implementierung des ganzen Beratungsinstruments gefährdet.

27.2.8 Sekundärprävention

Auch für den Bereich der Sekundärprävention möchte ARRIBA eine fundierte Prognose liefern, damit Patienten sich ein Bild von möglichen präventiven Effekten (Medikamente, Verhaltensänderungen) machen können.

Da es hier vor allem um die Demonstration von Präventions- bzw. Therapie-Effekten geht, ist der Mangel an „therapie-naiven“ Kohorten besonders gravierend [20]. Wir haben uns entschieden, die Kohorte einer aktuellen Medikamentenstudie an Patienten mit stabiler Angina pectoris (ACTION-Trial) zugrunde zu legen [21]. Diese gibt einen aktuellen Morbiditäts- und Behandlungsstandard wieder, außerdem finden sich bei Studien dieser Art minimale Drop-out-Raten und ein optimales Monitoring in Bezug auf relevante Ereignisse (hohe interne Validität). Man kann auch davon ausgehen, dass Prüfarzte bei placebokontrollierten Medikamentenstudien weniger selektiv rekrutieren als z.B. bei Studien zu perkutanen oder operativen Interventionen (bei letzteren stärker ausgeprägte Patienten- oder Arzt-Präferenzen), was die externe Validität erhöht.

Für das Zielkriterium Tod, MI oder Schlaganfall ergab sich hier in fünf Jahren Follow-up eine Wahrscheinlichkeit von 15%, die wir für zehn Jahre auf 30% extrapolieren haben. Da dies jedoch eine insgesamt eher intensiv behandelte Gruppe war, gehen wir von einer Relativen Risiko-Reduktion durch die verschiedenen Interventionen bei den Studienteilnehmern von insgesamt 40% aus. Mit anderen Worten: bei Nicht-Behandlung in Bezug auf Prognoseverbesserung wäre bei 50% der Patienten ein kardiovaskuläres Ereignis aufgetreten. Deshalb wird in der ARRIBA-Herz-Beratung für die Sekundärprävention (manifeste arteriosklerotische Erkrankung) vereinfachend von einer 10-Jahres-Ereignisrate von 50% ausgegangen.

27.2.9 Präventive Effekte

Nach der Berechnung einer individuellen Risikoprognose ist die Demonstration präventiver Effekte der nächste Schritt der Beratung nach ARRIBA-Herz. Diese sind für medikamentöse Maßnahmen so gut dokumentiert wie in kaum einem anderen Gebiet der Medizin. Nicht nur große randomisierte kontrollierte Studien mit mehrjährigem Follow-up von klinisch relevanten Outcomes, sondern auch systematische Übersichtsarbeiten mit Meta-Analysen unterstützen hier die Empfehlungen für Antihypertensiva [22; 23], Statine [24] und Thrombozyten-Aggregations-Hemmer [25].

Die Effekte von Verhaltensänderungen (Ernährungsumstellung, Nikotinkarenz, mehr Bewegung) sind dagegen nur durch Beobachtungsstudien gut belegt [26]. Hier besteht natürlich die Gefahr, dass trotz Adjustierung für Confounder in den Primärstudien die Wirkungen von Rauchstopp und vermehrter körperlicher Aktivität überschätzt werden. Für eine Ernährungsumstellung (Fettreduktion) liegen inzwischen Ergebnisse einer randomisierten kontrollierten Interventionsstudie mit klinisch relevanten Zielkriterien vor [27]; die Resultate sind ernüchternd, da sich keine Wirkung auf klinisch relevante Ereignisse zeigte.

Die mit ARRIBA-Herz arbeitenden Hausärzte sind für eine quantitative Darstellung der Effekte von Verhaltensumstellungen dankbar. Allerdings wird der motivierende Effekt einer solchen Demonstration oft überschätzt. Dazu sind weitergehende Gesprächstechniken erforderlich (z.B. motivierende Gesprächsführung) [28].

Die Rückmeldungen von Hausärzten legen es auch nahe, einen „Anker“ für die Beratung festzulegen, d.h. eine Risikoschwelle, oberhalb derer eine Intervention grundsätzlich sinnvoll ist. Die Statin-Studien mit der geringsten kardiovaskulären Ereignis-Inzidenz im Kontrollarm sind die Primärpräven-

tions-Studien AF-CAPS/Tex-CAPS [29] und WoS [30]: 7.7 in 5.2 J bzw. 7.9 in 5 Jahren. Diese lassen sich auf etwa 15% in 10 Jahren extrapolieren. Oberhalb dieser Risikoschwelle wissen wir also durch Studien, dass der Nutzen die unerwünschten Effekte überwiegt (Evidenzschwelle). Allerdings heißt dies nicht, dass oberhalb dieser Schwellen eine Intervention erfolgen „muss“. Im Einklang mit der Philosophie partizipativer Entscheidungsfindung beginnt hier erst die offene Diskussion mit dem Patienten.

27.2.10 Visuelle Darstellung

Um Verzerrungen und Missverständnisse zu vermeiden, stellen wir die 10-Jahres-Wahrscheinlichkeit für kardiale und zerebrovaskuläre Ereignisse als natürliche Häufigkeit anhand von 100 Smilies dar (s. Abb. 28.1) [31]. Dabei sind die beratenden Hausärzte angehalten, Ausdrücke wie „Prozent“, „Risiko“ o.Ä. zu vermeiden.

Zwar haben optische Charts mit farbiger Kennzeichnung für Wahrscheinlichkeitsbereiche (z.B. SCORE) den Vorteil einfacher Handhabung für den beratenden Arzt. Sie bieten jedoch für den Patienten nur eine geringe Anschaulichkeit. Dies ist der Grund, warum ARRIBA-Herz das Punktesystem mit der nachfolgenden Umrechnung in eine Wahrscheinlichkeit einsetzt.

27.2.11 Gesprächsführung

Was ARRIBA-Herz von sämtlichen anderen Kalkulationshilfen unterscheidet, ist die Einbettung der Risikoprognose in ein komplettes „Paket“ für die Beratung in der Praxis. Aus didaktischen Gründen haben wir das Vorgehen in Form von sechs Schritten definiert, die einem „Drehbuch“ entsprechen (s. Tab. 28.1). Da dies mit der Struktur anderer hausärztlicher Routinen übereinstimmt [32], kann es in

das alltägliche Vorgehen integriert und mit Zeit und Übung als kaum noch bewusste Routine „abgespeichert“ werden.

Neben einer möglichst präzisen Gesamtrisiko-Prognose ist die Einbeziehung des Patienten in die Entscheidung eine zweite zentrale Idee von ARRIBA-Herz. Je nach dem etablierten Praxisstil einzelner Hausärzte verlangt dies einen gewissen Mentalitätssprung. In Fortbildungsveranstaltungen wird dies gleichsam im Simulator geübt, wobei Teilnehmer und/oder Leiter ein strukturiertes Feedback geben, das sich an den sechs ARRIBA-Schritten orientiert. Für die ersten Einsätze in der Praxis wird eine entsprechende Selbstbeurteilungs-Hilfe bereitgestellt, die sich ebenfalls an den leicht zu merkenden sechs Schritten orientiert. In der Broschüre geben wir ausführliche Hilfen für eine Gesprächsführung einschließlich Formulierungshilfen.

Auch wenn eine kontrollierte Studie gezeigt hat, dass in ARRIBA-Herz geschulte Hausärzte ihre Patienten in höherem Maße in eine Entscheidung über Maßnahmen der kardiovaskulären Prävention einbeziehen, so muss man diesen Paradigmenwechsel weg vom patriarchalischen Arztbild in einer langfristigen Perspektive sehen. Obwohl der Effekt einer Fortbildung gemäß ARRIBA-Herz nur mäßig ausgeprägt gewesen ist, haben wir es sicher mit einer säkularen, tief und weit greifenden Entwicklung zu tun. Langfristig werden immer mehr Patienten bei einer immer größeren Palette von Entscheidungen valide und individualisiert beraten werden und diese Beratung auch einfordern.

Schließlich bietet die hausärztliche Beratungssituation die Chance, das unterschiedlich ausgeprägte Bedürfnis von Patienten zur Teilnahme individuell zu berücksichtigen. Erfahrungsgemäß schrecken viele Menschen vor einer aktiven Teilnahme an Entscheidungen zurück; in dieser Situation ermöglicht ARRIBA-Herz immerhin eine Risikokalkulation durch den Arzt.

ARRIBA-Herz 😊	Name: _____ Datum: _____ Arzt: _____	Beratungsbogen
----------------------	---	-----------------------

3.3 Risiko objektiv: Wahrscheinlichkeit optisch darstellen

Stellen Sie sich 100 Menschen vor, welche die gleichen Werte haben wie Sie. Von diesen werden in den nächsten 10 Jahren einen Herzinfarkt oder Schlaganfall erleiden oder daran sterben; werden gesund bleiben.

Wenn alle 100 „Doppelgänger“ 10 Jahre lang Vorbeuge-Maßnahmen durchführen, werden weitere nicht betroffen sein; trotz Prävention werden erkranken oder sterben.

Risikopunkte	😊 😊 😊 😊 😊 😊 😊 😊 😊 😊
Alter:	😊 😊 😊 😊 😊 😊 😊 😊 😊 😊
Rauchen:	😊 😊 😊 😊 😊 😊 😊 😊 😊 😊
Ges.-Chol.:	😊 😊 😊 😊 😊 😊 😊 😊 😊 😊
HDL-Chol.:	😊 😊 😊 😊 😊 😊 😊 😊 😊 😊
Blutdruck:	😊 😊 😊 😊 😊 😊 😊 😊 😊 😊
Summe:	😊 😊 😊 😊 😊 😊 😊 😊 😊 😊
Zusatzrisiken (Familienanamnese, Diabetes – siehe lami-n. Bogen)	😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊
10-Jahres-Risiko:	😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊 😊

4 Informationen über Möglichkeiten zur Vorbeugung (Prävention)

Vorbeuge-Maßnahmen	Relative Risikominderung	Individueller Plan
Rauch-Stopp	ca. 35%	
Bewegung 2–3x pro Woche > 30 Minuten	ca. 35%	
Fisch 2x pro Woche oder Omega-3-Fettsäuren	ca. 15%	
Blutdrucksenkung mit Medikamenten	ca. 25%	
ASS 100 mg pro Tag	ca. 20%	
Statin	ca. 20–25%	

5 Beiderseitige Bewertung der Möglichkeiten: Alternativen, Vor- und Nachteile

6 Absprache über weiteres Vorgehen: Verhaltensänderungen (Rauchen, Bewegung, Ernährung) und Medikamente. Was zuerst?

Termin Nachbesprechung?

Abb. 27.1: Beratungsbogen

27.2.12 Implementierung

Dass ARRIBA-Herz in zahlreichen Praxen so erfolgreich implementiert werden konnte, dürfte verschiedene Ursachen haben. Dazu zählt sicher der mehrjährige Entwicklungsprozess: das systematisch eingeholte Feedback der Anwender hat es uns möglich gemacht, das Instrument beständig zu optimieren und den Bedürfnissen der Praxis anzupassen [33].

Vor allem bei Medikamenten löst ARRIBA-Herz ein Dilemma, in dem sich verschreibende Hausärzte befinden: einerseits wird von Fachgesellschaften und finanziell oft nicht unabhängigen Meinungsführern eine immer aggressivere Behandlung von Stoffwechsellababweichungen gefordert; andererseits haben Ärzte sich an das Wirtschaftlichkeitsgebot zu halten. Ein großer Teil der Patienten nimmt selbst nach herkömmlicher „patriarchalischer“ Beratung verordnete Medikamente nicht ein bzw. setzt Verhaltensänderungen nicht um, scheint also vom Sinn der Maßnahmen nicht überzeugt zu sein. Auch in der offenen Beratung wie ARRIBA-Herz zeigen sich Patienten oft wenig beeindruckt von den Effekten gerade der medikamentösen Primärprävention, allerdings mit dem Unterschied, dass Medikamente, die der Patient ohnehin nicht nehmen würde, gar nicht erst verordnet werden. Dies entspricht anderen Untersuchungen zum Shared decision making; demnach führt die Einbeziehung des Patienten zu einem geringeren Einsatz von invasiven und kostenträchtigen Therapien [34; 35].

Vielleicht hat auch eine Rolle gespielt, dass eine **hausärztlich** entwickelte Entscheidungshilfe ein größeres Vertrauen bei den Anwendern genießt als ein Instrument, mit dem möglicherweise eine Partikular-Agenda verfolgt wird. So geht es ARRIBA-Herz ausdrücklich nicht darum, den Patienten zur Medikamenten-Einnahme zu bewegen. Vielmehr wird ein Erfolg der Beratung darin ge-

sehen, dass der Patient wirksam informiert wird und eine Entscheidung gemäß seinen Bedürfnissen und Prinzipien treffen kann.

Die Erfahrungen mit vielen hundert Anwendungen haben auch Fehlermöglichkeiten aufgezeigt; dazu gehören beispielsweise die Nicht-Berücksichtigung von manifesten arteriosklerotischen Erkrankungen und die Diabetes-Adjustierung der Risikopunkte statt des Risikos. Mühsam ist auch die Berechnung von präventiven Effekten aus dem kardiovaskulären Gesamtrisiko des Patienten und der Relativen Risiko-Reduktion einer Maßnahme. Dies wird weiter erschwert, wenn eine „Kette“ von mehreren Maßnahmen diskutiert werden muss, die doch recht komplexe arithmetische Operationen vom Berater verlangen. Diese Schwierigkeiten dürften durch die derzeit in Entwicklung befindliche elektronische Version von ARRIBA-Herz deutlich erleichtert werden.

27.3 Fazit

Die Verfügbarkeit wirksamer Behandlungen setzt verschreibende Ärzte unter Handlungsdruck. Sie müssen zwischen Hoch- und Niedrigrisiko-Patienten unterscheiden, um die vorhandenen Ressourcen patientengerecht sinnvoll und ökonomisch einzusetzen.

Dabei ist keines der vorhandenen Berechnungsinstrumente ideal [36]. Schon deshalb, weil es keine „therapie-naiven“ Kohorten gibt, werden wir auch in Zukunft nicht über bessere Instrumente als die oben diskutierten verfügen. Will man ein praxisfähiges Instrument entwickeln und verbreiten, sind deshalb Entscheidungen zwischen nicht-idealen Alternativen nötig, die zwangsläufig zum Hintanstellen bestimmter Aspekte führen.

Wir danken Prof. Dr. med. H.-W. Hense (Universität Münster) für hilfreiche Kommentare zur Entwicklung von ARRIBA-Herz.

Literatur

- [1] British Cardiac Society et al., Joint British recommendations on prevention of coronary heart disease in clinical practice: summary. British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, British Diabetic Association. *BMJ*. (2000), 320, 705–708
- [2] Krones T et al. (2006) Involving the patient in cardiovascular prevention: first results of a cluster randomized trial in Germany. *Wonca Europe ESPG/FM*, Florenz, 27.–30. August 2006. BMBF FKZ 01GK0401
- [3] Anderson KM et al., An updated coronary risk profile. A statement for health professionals. *Circulation*. (1991), 83, 356–362
- [4] Wilson PW et al., Prediction of coronary heart disease using risk factor categories. *Circulation*. (1998), 97, 1837–1847
- [5] Sullivan LM, Massaro JM, D’Agostino RB, Sr., Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med*. (2004), 23, 1631–1660
- [6] Assmann G, Cullen P, Schulte H, Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Munster (PROCAM) study. *Circulation*. (2002), 105, 310–315
- [7] Conroy RM et al., Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J*. (2003), 24, 987–1003
- [8] Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III), Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP). *JAMA*. (2001), 285, 2486–2497
- [9] Brindle P et al., Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study. *BMJ*. (2003), 327, 1267
- [10] Hense HW et al., Framingham risk function overestimates risk of coronary heart disease in men and women from Germany – results from the MONICA Augsburg and the PROCAM cohorts. *Eur Heart J*. (2003), 24, 937–945
- [11] Wolf PA et al., Probability of stroke: a risk profile from the Framingham Study. *Stroke*. (1991), 22, 312–318
- [12] Tunstall-Pedoe H, (ed.) (2003) MONICA Monograph and multimedia sourcebook. Coronary events table 7, Stroke events Table 6.1. WHO, Geneva
- [13] Gigerenzer G (2004) Das Einmaleins der Skepsis. Über den richtigen Umgang mit Zahlen und Risiken. Berliner Taschenbuch Verlag, Berlin
- [14] Haffner SM et al., Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction. *N Engl J Med*. (1998), 339, 229–234
- [15] Evans JM, Wang J, Morris AD, Comparison of cardiovascular risk between patients with type 2 diabetes and those who had had a myocardial infarction: cross sectional and cohort studies. *BMJ*. (2002), 324, 939–942
- [16] Lee CD et al., Cardiovascular events in diabetic and nondiabetic adults with or without history of myocardial infarction. *Circulation*. (2004), 109, 855–860
- [17] Vaccaro O et al., Impact of diabetes and previous myocardial infarction on long-term survival: 25-year mortality follow-up of primary screenees of the Multiple Risk Factor Intervention Trial. *Arch Intern Med*. (2004), 164, 1438–1443
- [18] Selvin E et al., Meta-analysis: glycosylated hemoglobin and cardiovascular disease in diabetes mellitus. *Ann Intern Med*. (2004), 141, 421–431
- [19] Stevens RJ et al., The UKPDS risk engine: a model for the risk of coronary heart disease in Type II diabetes (UKPDS 56). *Clin Sci (Lond)*. (2001), 101, 671–679
- [20] Hense HW, Observations, predictions and decisions – assessing cardiovascular risk assessment. *Int J Epidemiol*. (2004), 33, 235–239
- [21] Poole-Wilson PA et al., Effect of long-acting nifedipine on mortality and cardiovascular morbidity in patients with stable angina requiring treatment (ACTION trial): randomised controlled trial. *Lancet*. (2004), 364, 849–857
- [22] Turnbull F, Effects of different blood-pressure-lowering regimens on major cardiovascular events: results of prospectively-designed overviews of randomised trials. *Lancet*. (2003), 362, 1527–1535
- [23] Psaty BM et al., Health outcomes associated with various antihypertensive therapies used as first-line agents: a network

- meta-analysis. *JAMA*. (2003), 289, 2534–2544
- [24] Baigent C et al., Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet*. (2005), 366, 1267–1278
- [25] Antithrombotic Trialists' Collaboration, Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. *BMJ*. (2002), 324, 71–86
- [26] Peto R et al., Mortality from tobacco in developed countries: indirect estimation from national vital statistics. *Lancet*. (1992), 339, 1268–1278
- [27] Howard BV et al., Low-fat dietary pattern and risk of cardiovascular disease: the Women's Health Initiative Randomized Controlled Dietary Modification Trial. *JAMA*. (2006), 295, 655–666
- [28] Keller S, Velicer WF, Prochaska JO (1999) Das Transtheoretische Modell-Eine Übersicht. In: Keller S, Motivation zur Verhaltensänderung. Das Transtheoretische Modell in Forschung und Praxis, 17–44. Lambertus, Freiburg
- [29] Downs JR et al., Primary prevention of acute coronary events with lovastatin in men and women with average cholesterol levels: results of AFCAPS/TexCAPS. Air Force/Texas Coronary Atherosclerosis Prevention Study. *JAMA*. (1998), 279, 1615–1622
- [30] Shepherd J et al., Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med*. (1995), 333, 1301–1307
- [31] Gigerenzer G, Wie kommuniziert man Risiken? *Fortschr Fortbild Med*. (2002), 13–22
- [32] Hamm RM, Scheid DC, Smith WR, Tape TG (2003) Opportunities for applying psychological theory to improve medical decision making: Two case histories. In: Champman GB, Sonnenbert FA, Decision making in health care: theory, psychology, and applications, 386–421. Cambridge Univ. Pr., Cambridge
- [33] Sadowski EM et al., Evaluation komplexer Interventionen: Implementierung von AR-RIBA-Herz?, einer Beratungsstrategie für die Herz-Kreislaufprävention. *ZFA*. (2005), 81, 429–434
- [34] O'Connor AM et al., Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev*. (2003), CD001431
- [35] Kennedy AD et al., Effects of decision aids for menorrhagia on treatment choices, health outcomes, and costs: a randomized controlled trial. *JAMA*. (2002), 288, 2701–2708
- [36] Lenz M, Muhlhauser I, Kardiovaskuläre Risikoschätzung für eine informierte Patientenentscheidung. Wie valide sind die Prognoseinstrumente? *Med Klin (Munich)*. (2004), 99, 651–661

28a Revaskularisation bei koronarer Herzerkrankung

Martin Ruß, Karl Werdan

28a.1 Klinisches Szenario

Ein 72-jähriger Mann in gutem Allgemeinzustand wurde vom Hausarzt wegen typischer pektanginöser Beschwerden, die regelmäßig bei mittelschwerer Belastung aufgetreten waren, zunächst mit einem Nitrat, später zusätzlich, aufgrund einer gleichzeitig bestehenden arteriellen Hypertonie weiterhin mit einem Betablocker behandelt. Nach anfänglicher Besserung der Symptomatik konnten allerdings im weiteren Verlauf die Symptome für den Patienten nicht zufriedenstellend kontrolliert werden. Daher erfolgte die Überweisung des Patienten zum Kardiologen, der aufgrund eines pathologischen Belastungs-EKGs mit Ischämie-bedingten Senkungen in den Vorderwandableitungen die Indikation zur Herzkatheteruntersuchung stellte. Hierbei zeigt sich eine koronare Drei-Gefäßerkrankung mit einer proximalen 90%igen Stenosierung des Ramus interventrikularis anterior (RIVA) sowie ebenfalls hochgradigen Stenosierungen des medialen Ramus circumflexus (RCX) und der distalen rechten Koronararterie (RCA). Eine vorher durchgeführte Echokardiographie ergab eine normale linksventrikuläre Pumpfunktion.

Der konkrete Befund des Koronarstatus bildet den Ausgangspunkt für die Überlegungen zur evidenzbasierten optimalen Therapie des Patienten. Die koronare Herzerkrankung stellt in Bezug auf die optimale Therapie ein vielschichtiges Krankheitsbild dar. Einerseits kann die Symptomatik der stabilen Angina zur Einschränkung der Belastbarkeit bis hin zur absoluten Immobilität führen; des Weiteren besteht die Gefahr eines akuten Myokardinfarktes und darüber hinaus resultiert eine Stadien-abhängige Verschlechterung der Prognose *quoad vitam*. Der Pa-

tient erhofft sich von der Therapie eine Verbesserung seiner Beschwerdesymptomatik sowie die Verhinderung des krankheitsbedingt vorzeitigen Versterbens – und dies möglichst nebenwirkungs- und komplikationsarm. Aufgrund dieser Erwartungslage genügt eine Bewertung der Evidenz in Bezug auf die Überlebensrate nicht, um sowohl dem behandelnden Arzt als auch dem Patienten bei der Entscheidung für die jeweilige Therapieoption zu helfen. Daher wird die Evidenz der verschiedenen Therapieziele, soweit dies aufgrund der Studienlage möglich ist, getrennt bewertet.

28a.2 Revaskularisation oder medikamentöse Therapie?

Die spezifische medikamentöse Therapie der Angina pectoris bei unserem Patienten besteht in der Gabe eines Nitrats und eines Betablockers (bei Betablocker-Unverträglichkeit von Ivabradin); falls erforderlich, kann die antianginöse Behandlung um einen Kalzium-Antagonisten mit den bekannten Einschränkungen erweitert werden.

Im oben geschilderten Fallbeispiel einer koronaren Drei-Gefäßerkrankung kommen für die Revaskularisation operative (aortokoronare Bypassoperation/ACB) und interventionelle Techniken (perkutane Koronarangioplastie/PCI, meist mit Stent-Implantation in gleicher Sitzung) infrage. Allerdings konnten weder die interventionelle noch die operative Revaskularisationsmethode in aktuellen Studien eine Senkung der Mortalität im Vergleich zur medikamentösen Therapie erreichen. Dennoch empfiehlt die Nationale

Versorgungsleitlinie KHK [1] die operative Revaskularisation als primäres (starke Empfehlung ↑↑) und die interventionelle Revaskularisation als sekundäres Vorgehen (s. Abb. 28a.1). Die Evidenz, die dieser Empfehlung zugrunde liegt, ist jedoch nicht so eindeutig wie man glauben möchte, da die CASS-Studie bei Patienten mit koronarer Drei-

Gefäßkrankheit durch die ACB im Vergleich zur medikamentös-konservativen Therapie selbst nach acht Jahren keine signifikante Mortalitätsreduktion im Gesamtkollektiv der Patienten aufzeigen konnte. Lediglich in der Subgruppenanalyse der Patienten mit eingeschränkter linksventrikulärer Pumpfunktion ergab sich ein signifikanter Überlebensvor-

Koronare Herzkrankheit mit signifikanter ($\geq 50\%$) linkskoronarer Hauptstammstenose	
12-2	
<ul style="list-style-type: none"> • Bei linkskoronarer signifikanter Hauptstammstenose soll die operative Revaskularisation (ACB) angestrebt werden. Sie ist in Bezug auf Überleben, MACE und Lebensqualität der PCI und der konservativen Therapie überlegen. • Inoperablen Patienten und Patienten, die nach sorgfältiger Aufklärung eine operative Revaskularisation ablehnen, kann alternativ die PCI empfohlen werden. Dies gilt für die Therapieziele Verbesserung der Prognose und Lebensqualität. 	↑↑
Koronare Mehrgefäßkrankung mit hochgradigen proximalen Stenosen ($> 70\%$)	
12-3	
<ul style="list-style-type: none"> • Bei Patienten mit Mehrgefäßkrankungen sollen revaskularisierende Maßnahmen empfohlen werden, da dadurch die Lebensqualität erhöht werden kann und sie – nach Expertenmeinung und Registerdaten – auch zu einer Verbesserung der Prognose führen. • Bei Mehrgefäßkrankungen soll eine komplette Revaskularisation angestrebt werden. • Bei 3-Gefäßkrankung ist der ACB das primäre Vorgehen und die PCI das sekundäre Vorgehen. 	↑↑
12-4	
Patienten mit proximaler RIVA-Stenose ($\geq 70\%$) sollten unabhängig von der Symptomatik einer revaskularisierenden Maßnahme zugeführt werden.	↑
12-5	
Alle anderen Patienten ohne RIVA-Stenose mit symptomatischer, medikamentös nicht adäquat beherrschbarer Eingefäßkrankung sollen mit einer revaskularisierenden Maßnahme (in der Regel PCI) aus antianginöser Indikation behandelt werden.	↑↑
12-6	
<ul style="list-style-type: none"> • Älteren Patienten (> 75 Jahre) mit ausgeprägter, persistierender, trotz medikamentöser Therapie bestehender Symptomatik soll die Revaskularisation empfohlen werden. • PCI und ACB führen im Vergleich zur medikamentösen Therapie zu einer deutlichen symptomatischen Verbesserung der KHK, ohne eine erhöhte Sterblichkeit zu bedingen. Sie sollten auch bei alten Patienten mit ausgeprägter persistierender Symptomatik trotz medikamentöser Therapie empfohlen werden. 	↑↑

Abb. 28a.1: NVL-KHK: Empfehlungen zur Revaskularisations-Therapie

teil gegenüber der ausschließlich medikamentös behandelten Patientengruppe [2]. Hierbei muss allerdings bedacht werden, dass diese Studie 1977 begonnen worden war, also zu einer Zeit, in der ACE-Hemmer nicht als Standard der Herzinsuffizienz-Therapie galten, Betablocker in der Herzinsuffizienz-Therapie kontraindiziert und Statine noch nicht verfügbar waren.

Beide Revaskularisationsmethoden führen jedoch im Vergleich zur medikamentösen Therapie zu einer deutlichen symptomatischen Verbesserung (Reduktion der Angina pectoris) sowie zu einer Verbesserung der Lebensqualität (Quality of Life), wobei dieser Effekt in manchen Studien bei der operativen Revaskularisation noch etwas ausgeprägter ist [3].

Die Komplikationsrate, d.h. das Überleben ohne weitere schwerwiegende kardiovaskuläre Ereignisse (MACE/major adverse cardiovascular events) – je nach Studie definiert als kombinierter Endpunkt aus Tod, nicht-letalem Myokardinfarkt, Schlaganfall oder erneuter Revaskularisation – ist bei der operativen Revaskularisation am günstigsten; medikamentöse Therapie und interventionelle Revaskularisation liegen hier auf vergleichbarem Niveau, wobei jedoch beachtet werden muss, dass dies im Falle der PCI nahezu ausschließlich durch die höhere Rate an erneut notwendigen Revaskularisationen aufgrund einer in-Stent-Restenose bedingt ist.

28a.3 Klinisches Szenario

Aufgrund der Datenlage können dem Patienten mit koronarer Drei-Gefäßkrankung und normaler linksventrikulärer Pumpfunktion alle drei Möglichkeiten, nämlich ein medikamentöser Therapieversuch ebenso wie die operative und die interventionelle Revaskularisation angeboten werden. Gleichzeitig sollte dem Patienten auch mitgeteilt werden, dass eine Revaskularisation

mit einer deutlichen Verbesserung der Symptomatik und Lebensqualität sowohl interventionell als auch operativ durchgeführt werden kann, ohne eine höhere Sterblichkeit zu bedingen, d.h. das operative bzw. interventionelle Risiko wird im Langzeitverlauf durch die Revaskularisation vollständig kompensiert.

Der Patient ist durch die koronare Drei-Gefäßkrankung in seinem Aktionsradius trotz der Behandlung mit zwei unterschiedlichen antiaanginösen Präparaten (Nitrat und Betablocker) stark eingeschränkt. Die Möglichkeit einer Verbesserung der Symptomatik durch Intensivierung der Therapie mit einem Ca²⁺-Antagonisten wird vom behandelnden Arzt mit dem Patienten besprochen, aber insgesamt als unwahrscheinlich in Bezug auf eine deutliche Verbesserung der Symptomatik eingeschätzt. Zudem lehnt der Patient eine weitere Steigerung seiner täglichen Medikation ab, sodass er sich für eine koronare Revaskularisation entscheidet und den Arzt um weitere Ratschläge bezüglich der Art der Revaskularisation bittet.

28a.4 Revaskularisation: operativ oder interventionell?

28a.4.1 Einfluss der Revaskularisations-Methode auf die Prognose

Aufgrund der eher guten Prognose der stabilen Angina pectoris bei erhaltener linksventrikulärer Pumpfunktion ergaben die Letalitätsanalysen meist nur sehr geringe Unterschiede zwischen den einzelnen Therapiekonzepten. In mehreren randomisierten Studien wurden die koronarchirurgische und die interventionelle Revaskularisation bei koronarer Mehrgefäßkrankung verglichen. Darüber hinaus gibt es eine Reihe von Registerarbeiten, die diese Therapieoptionen vergleichend gegenüberstellen. Ein grundlegender Unterschied von Studien und Registerarbeiten besteht darin, dass in den randomisierten Studien meist nur ein kleiner An-

teil der gescreenten Patienten eingeschlossen worden ist. Dies ist u.a. dadurch bedingt, dass die Prüffärzte nur diejenigen Patienten einschlossen, von denen sie annahmen, dass sie durch beide Techniken adäquat behandelt werden könnten. Lässt der Koronarstatus bei Patienten ein erfolgreiches interventionelles Vorgehen erwarten, sollte primär eine operative Lösung angeboten werden. Andererseits bietet sich bei beträchtlicher Komorbidität und hohem operativem Risiko, wie z.B. bei schwerer Lebererkrankung oder schwerer chronisch-obstruktiver Lungenerkrankung mit hohem anästhesiologischem und vor allem postoperativem Risiko, eine interventionelle Revaskularisation an.

Die BARI-Studie (Bypass Angioplasty Revascularisation Investigation) konnte nach fünf Jahren einen Trend [4] und nach sieben Jahren eine signifikant verringerte Mortalität nach Bypass-Operation im Vergleich zur PCI [5] nachweisen. Die Analyse der Subgruppen ergab, dass der Vorteil bezüglich der operativen Revaskularisation ausschließlich auf das verbesserte Überleben in der Gruppe der Diabetiker zurückzuführen war. Keine Unterschiede ergaben sich in der Gruppe der Nicht-Diabetiker, weder bezüglich der Anzahl der erkrankten Gefäße (Zwei- oder Drei-Gefäß-Erkrankung) noch der Beteiligung des proximalen R. interventricularis anterior (RIVA vor Abgang des 1. Diagonalastes). Ähnliche Ergebnisse wie die der BARI-Studie wurden in mehreren weiteren Studien, wie RITA, MASS-II, GABI, ARTS, ERAC-I, EAST, CABRI u.a. bestätigt. Davon abweichende Ergebnisse erbrachten lediglich die SoS-Studie [Surgery or Stent], die ein besseres Resultat für die CABG-Gruppe ergab, und die ERAC-II-Studie [Argentinische randomisierte Studie PTCA vs. CABG], die ein besseres Ergebnis für die interventionelle Revaskularisation geliefert hatte. Diese Ergebnisse hatten sich vor allem ergeben, weil die SoS-Studie eine im Vergleich zu den anderen Studien sehr niedrige Mortalität in der CABG-Gruppe er-

bracht hatte, während die ERAC-II Studie eine vergleichsweise hohe Sterblichkeit im Zweig der operativen Revaskularisation hatte und so ein signifikant besseres Ergebnis zugunsten der PTCA-Gruppe resultierte. Eine Meta-Analyse der bis 2003 durchgeführten Studien ergab einen Vorteil für die Bypassoperation bezüglich der Mortalität in der Subgruppenanalyse der Patienten nur nach 5 und 8 Jahren, wobei in den Studien, welche 5- und 8-Jahres-Überlebensdaten erbrachten, keine Stents verwendet worden waren (als Endpunkte war die Mortalität nach 1, 3, 5 und 8 Jahren untersucht worden, bei Patienten mit Diabetes mellitus zusätzlich die 4-, 5- und 6-Jahres-Mortalität) [6]. Der Trend zu einem besseren Überleben zugunsten der Bypassoperation nach ein und drei Jahren in der Prä-Stent-Ära war nach Einführung der Stent-Implantation nicht mehr vorhanden, obwohl in diesem Zeitraum auch die Mortalität der Bypassoperation abgenommen hatte. Ebenso war der bei Diabetikern nach vier und fünf Jahren vorhandene Überlebensvorteil bei den Operierten nach sechs Jahren nicht mehr nachweisbar. Insgesamt scheinen die Unterschiede zwischen der operativen und der interventionellen Therapie in Bezug auf die Mortalität sehr gering zu sein; lediglich in der mittelfristigen Prognose – im Zeitraum von 5 Jahren nach Revaskularisation – scheint ein geringer Vorteil zugunsten der Bypassoperation zu bestehen, der sich vermutlich auf den Überlebensvorteil bei Patienten mit Diabetes mellitus stützt. Ob sich diese Daten auch für die Langzeitverläufe mit Stent-Implantation bestätigen, bleibt abzuwarten.

28a.4.2 Einfluss der Revaskularisations-Methode auf Symptomatik und MACE

Für den Patienten sind neben dem Endpunkt „Mortalität“ die Endpunkte „Verbesserung

der Symptomatik“ bzw. „Lebensqualität“ sowie die „MACE-Rate“ von ebenso großer Bedeutung.

Sowohl die interventionelle als auch die operative Revaskularisation reduzieren die Angina pectoris-Symptomatik stärker als die medikamentöse Therapie, wobei in einigen Studien die Symptomreduktion durch die operative Revaskularisation ausgeprägter ist als durch die Intervention. Die Bypass-Operation führt zu einer sehr guten mittel- und langfristigen Besserung der Angina: die Notwendigkeit für ein erneutes Revaskularisations-Verfahren – sei es erneute Bypassoperation oder PCI – ist gering. In diesen Punkten schneidet die operative Revaskularisation signifikant besser ab als die PCI. Dies liegt v.a. am Restenose-Problem der PCI. Durch die Einführung der Stent-Implantation konnte die Notwendigkeit zu erneuten Revaskularisationen nach ein und drei Jahren von 30% auf 15–20% reduziert werden [7]; die Einführung der Drug eluting stents lässt ein weiteres Absinken der Ereignisrate in Bezug auf die Revaskularisation von Restenosen erwarten. Bei Auftreten einer Restenose reicht jedoch meist eine erneute Intervention aus, um ein stabiles Langzeitergebnis zu erreichen. So konnte die GABI-Studie (German Angioplasty Bypass Surgery Investigation) mit der bisher längsten Verlaufsbeobachtung zeigen, dass nach 13 Jahren zwar die Zeitspanne bis zur ersten erneuten Revaskularisations-Maßnahme in der PCI Gruppe signifikant kürzer war, es sich insgesamt aber weder für die Anzahl der erneuten Revaskularisations-Maßnahmen noch für das Crossover zwischen Bypassoperation und PTCA signifikante Unterschiede ergaben [8]. Alle Vergleichsstudien zeigen, dass die erhöhte MACE-Rate der PCI ganz überwiegend auf die höhere Rate erneuter Revaskularisationen zurückzuführen ist; bei den anderen Kriterien ergaben sich allenfalls nur geringe Unterschiede. In den Analysen der MACE-Raten sieht man dementsprechend eine Divergenz

der Kurven in den ersten sechs Monaten, danach verlaufen die MACE-Kurven in nahezu allen Studien parallel, was die Resultate der GABI-Studie bestätigt.

28a.5 Klinisches Szenario

Nach der Darstellung der unterschiedlichen Möglichkeiten zur Koronarrevaskularisation tendiert der Patient aufgrund der guten symptomatischen Verbesserung und der geringen MACE-Rate der operativen Revaskularisation dazu, sich einer Bypass-Operation zu unterziehen. Er fragt daraufhin den Kardiologen nach seinem persönlichen operativen Risiko. Anhand des EuroScores (s. Abb. 28a.2) ergibt sich ein perioperatives Risiko von 2,06% für die Bypassoperation. Der Patient sollte darauf hingewiesen werden, dass zusätzlich zur Revaskularisation eine unterstützende medikamentöse Therapie für ein optimales Langzeitergebnis notwendig ist. Diese besteht derzeit aus einer Hemmung der Thrombozytenaggregation, der Sekundärprophylaxe durch Gabe eines Statins und der Fortsetzung der Beta-blocker-Therapie bzw. – bei Betablockerkontraindikationen – der Gabe von Ivabradin.

28a.6 Fazit

Die verschiedenen Therapieoptionen bei der stabilen koronaren Herzerkrankung sind in Bezug auf die Prognose gleichwertig, wenn eine normale linksventrikuläre Pumpfunktion und keine schwerwiegenden Begleiterkrankungen, v.a. ein Diabetes mellitus, vorliegen. Es ist daher gerechtfertigt, zunächst einen medikamentös-konservativen Therapieversuch zu unternehmen. In Bezug auf die Verbesserung der Symptomatik und der Lebensqualität zeigen sich deutliche Vorteile der beiden Revaskularisations-Therapien, sodass aufgrund dieser Ergebnisse auch eine primäre Revaskularisation angestrebt werden kann, insbesondere, wenn der Patient durch


	Additive EuroSCORE –	Logistic EuroSCORE $\beta_i \chi_i$
Patient-Faktoren		
Alter 72yr	3	0,928514094
Geschlecht <input type="checkbox"/> Female		
Chronische Lungenerkrankung <input type="checkbox"/> Yes		
Extrakardiale Gefäßerkrankung <input type="checkbox"/> Yes		
Neurologische Dysfunktion <input type="checkbox"/> Yes		
Vorherige Herzoperation <input type="checkbox"/> Yes		
Serum Kreatinine > 200 $\mu\text{mol/L}$ <input type="checkbox"/> Yes		
Aktive Endokarditis <input type="checkbox"/> Yes		
Kritischer präoperativer Zustand <input type="checkbox"/> Yes		
Kardiale Faktoren		
Instabile Angina pectoris <input type="checkbox"/> Yes		
Moderate LV-Dysfunktion (LVEF 30–50%) <input type="checkbox"/> Moderate		
Schlechte LV-Funktion (LVEF < 30%) <input type="checkbox"/> Poor		
Kürzlicher Myokardinfarkt <input type="checkbox"/> Yes		
Pulmonale Hypertension <input type="checkbox"/> yes		
Operative Faktoren		
Notfalloperation <input type="checkbox"/> Yes		
Weitere Operation am Herzen (neben CABG) <input type="checkbox"/> Yes		
Eingriff an der thorakalen Aorta <input type="checkbox"/> Yes		
Ventrikel Septum Ruptur <input type="checkbox"/> Yes		
 EuroSCORE	–	$e^{(-4.789594 + \sum \beta_i \chi_i)} / 1 + e^{(-4.789594 + \sum \beta_i \chi_i)}$
Downloaded from http://euroscore.org	3	2,06%

Abb. 28a.2: Berechnung des Euroscores zur Abschätzung des operativen Risikos

die Angina pectoris in seinem Aktionsradius eingeschränkt ist. Hierbei besteht sowohl die Möglichkeit einer operativen als auch einer interventionellen Revaskularisation. Bei ein-

geschränkter linksventrikulärer Pumpfunktion oder gleichzeitig bestehendem Diabetes mellitus sollte ebenfalls primär eine Revaskularisation angestrebt werden.

Literatur

- [1] Bundesärztekammer (BÄK), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Kassenärztliche Bundesvereinigung (KBV), Ärztliches Zentrum für Qualität in der Medizin (ÄZQ). Nationale Versorgungsleitlinie Chronische Koronare Herzkrankheit. <http://www.versorgungsleitlinien.de/themen/khk> (08.06.06)
- [2] Passamani E et al., A randomized trial of coronary artery bypass surgery. Survival of patients with a low ejection fraction. *N Engl J Med.* (1985), 312, 1665–1671
- [3] Favarato ME et al., Quality of life in patients with symptomatic multivessel coronary artery disease: A comparative post hoc analyses of medical, angioplasty or surgical strategies-MASS II trial. *Int J Cardiol.* (2006)
- [4] The Bypass Angioplasty Revascularization Investigation (BARI) Investigators, Comparison of coronary bypass surgery with angioplasty in patients with multivessel disease. *N Engl J Med.* (1996), 335, 217–225
- [5] The Bypass Angioplasty Revascularization Investigation (BARI) Investigators, Seven-year outcome in the Bypass Angioplasty Revascularization Investigation (BARI) by treatment and diabetic status. *J Am Coll Cardiol.* (2000), 35, 1122–1129
- [6] Hoffman SN et al., A meta-analysis of randomized controlled trials comparing coronary artery bypass graft with percutaneous transluminal coronary angioplasty: one-to eight-year outcomes. *J Am Coll Cardiol.* (2003), 41, 1293–1304
- [7] Hueb WA et al., Five-year follow-up of the medicine, angioplasty, or surgery study (MASS): A prospective, randomized trial of medical therapy, balloon angioplasty, or bypass surgery for single proximal left anterior descending coronary artery stenosis. *Circulation.* (1999), 100, II107–II113
- [8] Kaehler J et al., 13-year follow-up of the German angioplasty bypass surgery investigation. *Eur Heart J.* (2005), 26, 2148–2153

28b Chirurgische Revaskularisation bei koronarer Herzkrankheit

Robert von Wattenwyl, Hans-Reinhard Zerkowski

28b.1 Klinisches Szenario – Chirurgisches Konsilium

Klares Ziel jedes therapeutischen Ansatzes muss das Bestreben sein, das für den Patienten beste Behandlungsverfahren zu wählen. Die Entscheidung soll auf einer ausführlichen Aufklärung bezüglich aller drei Therapieoptionen der koronaren Herzkrankheit (medikamentös, interventionell, chirurgisch) im Hinblick auf Therapieerfolg, Symptomatik/Lebensqualität und Risiken basieren. „Gatekeeper“-Situationen, welche zu einer Bevorteilung (Bias) eines therapeutischen Verfahrens ohne Evidenz führen könnten, sollten vermieden werden. Dazu eignet sich neben der leitlinienbasierten Einführung von Behandlungspfaden besonders eine interdisziplinäre Fallbesprechung aller nicht nach Leitlinien eindeutig zu entscheidenden Situationen oder, falls dies mitunter nicht organisierbar ist, ein auf Leitlinien¹ gestütztes Konsilium, das die besonderen individuellen Bedingungen und Umständen des Patienten berücksichtigt.

Dem Konsiliarius präsentiert sich ein 72-jähriger Mann in gutem Allgemeinzustand mit arterieller Hypertonie, mittels beta-Adrenozep-tor-Blocker behandelt. Die nicht zufriedenstellend behandelbare Angina führte zur Herzkatherunter-suchung: Diagnostiziert wurde koronare Drei-Gefäß-Erkrankung mit proximalen Stenosen von RIVA, RCX und rechter Kranzarterie bei normaler Kammerfunktion. Von chirurgischer Revaskularisation sind in dieser Konstellation keine technischen Probleme zu erwarten.

Bei koronarer Drei-Gefäß-Erkrankung ist die operative Revaskularisation nach internationa-

lem Konsens [1] (AHA/ACC) die Methode der Wahl und nach Nationaler Versorgungsleitlinie das primäre Vorgehen vor PCI.

Welche Fakten und Überlegungen sind für den Patienten hinsichtlich Prognose, Symptomatik und MACE der drei prinzipiellen Therapieoptionen (und ihrer Weiterentwicklungen) wesentlich, um die Empfehlung nachvollziehen zu können? Da der Patient sich neben der Verhinderung vorzeitigen Sterbens eine Verbesserung der Symptomatik erhofft und von nebenwirkungsarmer und möglichst komplikationsfreier Behandlung ausgeht, sind die letzteren sekundären Parameter ebenfalls zu berücksichtigen.

28b.2 Bypassoperation oder medikamentöse Therapie?

Neuere prospektive Vergleichsstudien zwischen Operation und zeitgemäßer medikamentöser Therapie gibt es nicht; diese Frage stellte sich in den späten 60er-Jahren und Anfang der 70er zur Rechtfertigung der damals aufkommenden und sicher zum damaligen Zeitpunkt nicht ganz risikolosen Koronarchirurgie (in ihrer überwiegenden Ausführung als aorto-koronarer Venenbypass).

¹ Mit den Nationalen Versorgungsleitlinien existiert nun ein Instrument, mit welchem eine zielorientierte, gewissenhafte, stufengerechte und risiko-adaptierte Therapiewahl, evidenzbasiert und für Fachleute wie Laienöffentlichkeit nachvollziehbar, erleichtert werden soll und nach aktuell bester verfügbarer Evidenz die nachhaltig beste Therapieentscheidung des Arztes mit seinem Patienten unterstützt wird [MLS, Therapie nach Leitlinie kann Leben retten, Frankfurter Allgemeine Zeitung. 24.05.2006; <http://www.versorgungsleitlinien.de/themen/khk>].

Alle zu diesem Zweck durchgeführten Untersuchungen ergaben einen signifikanten Überlebensvorteil für die operierten Patienten [2,3,4], der für die CASS-Studie bei Patienten mit Drei-Gefäß-Erkrankung und normaler Ventrikelfunktion auch nach 16 Jahren noch besteht [5] – und das, obwohl auf der einen Seite im Verlauf der Zeit die verbesserte medikamentöse Therapie vermutlich zu einer kontinuierlichen Anpassung der Regime führte, während auf der anderen Seite die operierten Patienten ohne erneute Operation nicht vom Fortschritt der Chirurgie profitierten (obwohl schon Loop et al vor mehr als 2 Jahrzehnten beweisen konnten, dass gerade die Einführung der arteriellen Revaskularisation allein des RIVA mit der A. thoracica int. (vulgo IMA) zu signifikanter Verbesserung des Überlebens bei Mehr-Gefäß-Erkrankung führt [6]).

Danach gibt es lediglich eine einzige Studie, die als Nebenbefund wieder die Überlegenheit der Chirurgie gegenüber moderner medikamentöser Therapie – und erstmalig, unter bestimmten Konstellationen, auch die Überlegenheit von PCI bei niedergradigerer KHK – belegt; in dieser Studie zeigen Smith et al. (Duke, Durham), die die mehr als 18000 konsekutiv behandelten Patienten ihrer Institution in verschiedenen Ären über fast 2 Jahrzehnte bis in die Jetzt-Zeit analysiert haben, dass bei sorgfältiger Zuweisung der Patienten zur Koronarchirurgie schon nach wenigen Jahren ein statistischer Überlebensvorteil besteht [7].

28b.3 Klinisches Szenario – Nachfrage des Patienten

Der geistig rege, durch das Internet teilinformierte und wissbegierige Patient hakt nach; er habe gehört, die viel weniger belastende Katheterdehnung mit Einbringen eines Stents sei aber bei seiner Erkrankung prinzipiell gleichwertig. Dass Medikamente unterlegen seien, habe er ja selbst

bemerkt, da die Symptome immer wieder kämen. Er wolle die beste und sicherste Lösung, habe aber Angst vor zu hohen Risiken.

28b.4 Revaskularisation: Operation oder Katheterintervention? Hat die Operation Einfluss auf die Prognose?

Bei unserem Patienten geht es um die Konstellation einer koronaren Drei-Gefäß-Erkrankung mit proximalen Stenosen aller drei Hauptäste.

Prospektiv randomisierte Vergleichsstudien gehen, um keinen Patienten aktiv durch eine Studienteilnahme zu gefährden, aber immer von der Grundvoraussetzung aus, dass ein Patient nur an der Studie teilnehmen darf, wenn die Prüfarzte der gemeinsamen Überzeugung sind, dass der individuell zur Diskussion stehende Patient mit beiden Therapieoptionen (Operation oder Katheterintervention) gleich erfolgreich behandelt werden kann.

Brett et al [8] analysierten 12 randomisierte kontrollierte Studien (randomized controlled Trials, RCTs), welche PCI und CABG miteinander verglichen haben. Es zeigte sich, dass bezüglich des Überlebens zwischen den beiden Gruppen grundsätzlich keine Unterschiede bestehen. Bei genauerer Betrachtung fällt jedoch auf, dass bei längerem Follow-up, insbesondere in Subgruppen, ein Überlebensvorteil zugunsten von CABG sehr wohl deutlich wird. Dies wird in der Analyse von Hoffmann et al [9] deutlich, in welcher bezüglich Tod und fatalem Myokardinfarkt nach 1 und 3 Jahren kein Unterschied zwischen CABG und PCI bestand, jedoch nach 5 und 8 Jahren bezüglich kardialen Tod ein signifikanter Vorteil zugunsten der CABG-Gruppe dokumentiert werden konnte.

Aber auch kürzere Zeitintervalle demaskieren in dieser Analyse Unterschiede zu-

gunsten der CABG-Gruppe wie signifikant geringere Reinterventionsrate und deutlich geringeres Wiederauftreten von Angina (zusammen subsumiert als MACE). In der PCI-Gruppe hingegen zeigt sich eine deutlich höhere Re-Interventionsrate, wie sie beispielsweise schon in der MASS-II-Studie [10] beschrieben worden ist.

Noch schwerer wiegen diese Resultate, wenn man berücksichtigt, dass es sich gerade bei diesen (Vergleichs-) Studien mit anscheinend vergleichbaren Ergebnissen zwischen CABG und PCI bei den eingeschlossenen Patienten um ein hoch selektiertes Patientengut handelt.

Die in der Summe von Brett et al untersuchten Studien erfassen nach Analyse der Ein- und Ausschlusskriterien nur gerade einmal 3,4% des ursprünglich infrage kommenden Patientenguts mit Mehr-Gefäß-Erkrankung. Ausserdem werden oft in Vergleichsstudien trotz dieser Kriterien deutlich kränkere Patienten (schlechtere EF, mehr Komorbiditäten, höherer Anteil an Diabetikern) in der CABG-Gruppe mit gesünderen Patienten in der PCI-Gruppe verglichen.

Die Ergebnisse klassischer RCTs sind also auf den uns zum Konsil vorgestellten Patienten überhaupt nicht anwendbar².

Auf welches Erkenntnisniveau kann man sich dann zurückziehen? Statistisch und erkenntnistheoretisch nicht schlechter sind – mathematisch korrekt angewendet – Propensity-Score- und multivariante Analysen [11], die dann genau das aus RCTs ausgeschlossene Patientenkontinuum mit Drei-Gefäß-Erkrankung umfassen müssen und hohe Anforderungen an Register und Datenbanken stellen [12].

Propensity-score- (oder multivariante) Analysen wie die Studie von Brener [13] oder von Hannan [14] versuchen, die RCT-immanenten Nachteile durch die mathematisch korrekte Analyse dafür umso grösserer Register von konsekutiv behandelten Patientenserien im klinischen Alltag zu vermeiden und

belegen die Wichtigkeit der risikoadjustierten Analyse. Beiden Studien hielten nach Risikoadjustierung einen signifikanten Überlebensvorteil zugunsten der CABG-Gruppe bei Mehr-Gefäß-Erkrankungen fest. In der Studie von Hannan konnte sogar bei Drei-Gefäß-Erkrankung mit Beteiligung der proximalen LAD ohne Risikoadjustierung ein signifikanter Vorteil zugunsten CABG festgehalten werden.

28b.5 Revaskularisation: Operation oder Katheterintervention? Ist bei immer besserem Material die wiederholte Katheterbehandlung nicht auch eine Option anstelle Operation?

Vorzeitiges Versterben ist definitiv das härteste Kriterium zur Bestimmung der Überlegenheit einer Methode; die Verknüpfung zwischen Re-Stenose und Versterben ist jedoch nicht zwingend gegeben. Der Patient kann den letalen Infarkt auch durch Thrombose eines anderen Gefäßes als des zuvor mit Dehnung und Stent behandelten erleiden.

Die Erwartung der letzten Jahre, dass eine Verbesserung der Stent-Technologie (in nomine medikamentenbeschichtete Stents) zu einer Abnahme der Ereignisrate führt, darf aus zwei Gründen kritisch hinterfragt wer-

2 Zur Erklärung sei ein Editorial aus der Gruppe Zerkowski in JTCS von 2006 empfohlen: Quintessenz: Wenn nun einerseits gerade diejenigen Patienten, welche als interventionell nicht angebar gelten, von der Studie ausgeschlossen werden, andererseits diese nicht in die Studie eingeschlossenen Patienten aber tatsächlich gleichzeitig die typische Population der täglich im Alltag zu behandelnden („real world“) darstellen, entsteht zwangsläufig eine Deformierung der Wirklichkeit, sodass die Umsetzung der Studienergebnisse in den klinischen Alltag eine oftmals unzulässige Analogie darstellt (Grapow MT, von Wattenwyl R, Guller U, Beyersdorf F, Zerkowski HR. Randomized controlled trials do not reflect reality: real-world analyses are critical for treatment guidelines! J Thorac Cardiovasc Surg 2006. 132(1): 5-7.)

den: Erstens ist eine Re-Stenose kein klinischer Endpunkt, und zweitens führt die Anwendung intraluminal platzierter Depots mit zytotoxischen Medikamenten offenbar zu neuen gefäßschädigenden Phänomenen, deren Auswirkung unklar sind, aber in den bisher sehr kurze Zeitspannen umfassenden Studien unterschätzt wurden. Bis jetzt gibt es lediglich Befunde, die die Reduktion der Re-Stenoserate in der Frühphase nach Implantation belegen.

Die Studie von Bapabulle [15], welche den Benefit bezüglich der Mortalität von DES gegenüber BMS untersuchte, stellte fest, dass es keine einzige Studie gibt, welche zeigt, dass DES bezüglich der Mortalität überlegen ist. Auch der direkte Vergleich von DES und CABG im selben Zentrum lässt Skepsis aufkommen. So zeigte die Studie von Domburg et al [16], welche das Patientenprofil und Behandlungsdaten sowie Outcome von PCI und CABG vor und nach Einführung der DES verglich, trotz deutlich höheren Anteils an Drei-Gefäß-Erkrankungen (75% vs. 22%), Diabetes (24% vs. 17%) und eingeschränkter linksventrikulärer Pumpfunktion in der CABG-Gruppe der DES-Ära, ein besseres Outcome in der operativen Koronarrevaskularisation mit einer Mortalität von 2,2% (CABG) vs. 4.1% (PCI) nach 12 Monaten. Guyton [17] stützt dies und analysiert zusätzlich auch den Gesichtspunkt der Kosteneffektivität von CABG und Stent bei Mehr-Gefäß-Erkrankungen und hält nach 5 Jahren einen klaren ökonomischen Vorteil von CABG gegenüber Stent (BM und DES) fest.

Aktuelle Befunde zeigen darüber hinaus, dass unter Alltagsbedingungen bei Verwendung von DES bei Mehr-Gefäß-Erkrankungen bei Beendigung der antikoagulativen Medikation ein erhebliches Risiko letaler Spätthrombosierung eintritt [18].

28b.6 Klinisches Szenario – Alter und Risiko

Der Patient weist darauf hin, dass er sich zwar beileibe nicht alt fühle, auch sonst keine chronische Leiden habe, sich aber doch frage, wie Alter und Operation sich gegenseitig beeinflussten und wie das Risiko bei ihm infolge einer Bypassoperation einzuschätzen sei.

28b.7 Alter und Risikoabschätzung bei chirurgischer Koronarrevaskularisation

Gerade ältere Patienten profitieren maßgeblich von einer Revaskularisation. Fast alle Studien haben aber in der Vergangenheit Patienten über 70 Jahre ausgeschlossen. So gibt es nur grössere Register, die herangezogen werden können. Auch hier ist der invasive anmutende operative Eingriff mit zunehmendem Patientenalter gegenüber der PCI keinesfalls ein risikoreicherer Eingriff, wie die Studie von Graham [19] zeigt. Die Studie untersuchte Patientengruppen über und unter 70 Jahren, welche entweder medikamentös, mittels PCI oder mit CABG behandelt wurden. Es konnte eine signifikante Risikominderung bezüglich des Überlebens in der CABG-Gruppe gegenüber den anderen 2 Gruppen gezeigt werden; wobei das Resultat mit zunehmendem Alter deutlicher wurde.

Die Bypassoperation kann auch im höheren Lebensalter mit niedrigem Risiko angeboten werden [20]; die Sterblichkeit isolierter Bypasschirurgie liegt heute in Deutschland unter Einschluss aller Notfälle, Patienten im Schock, jeden Alters, etc. knapp unter 3%, unter 2% im unselektionierten Patientengut sind mehrheitlich zu beobachten. Im elektiven (geplanten Fall) liegt die Sterblichkeit heute deutlich unter 2, oft unter 1%. Der logistische Euroscore als griffige Grösse ist allerdings, wie unter <http://www.euroscore.org> erläutert, zur Risikoabschätzung im Hochrisiko-

obereich berechnet. Der lineare Score würde die Wirklichkeit eher treffen, greift aber bei unserem Patienten wieder zu kurz, da er nur einen Risikofaktor (Alter) aufweist. Dies ist allerdings heute der am wenigsten eingreifende, geht rechnerisch unabhängig von seiner Bedeutung ein. So hätte unser Patient bei Eingabe von 72, 68, 65 oder 60 Jahren ein rechnerisches Risiko von 2,07, 1,59, 1,31 und 0,94%.

Auf der anderen Seite muss dem Patienten auch mitgeteilt werden, dass er bei Stent-Behandlung einer Mehr-Gefäß-Erkrankung mindestens für 1 Jahr (eher, wie es jetzt schon scheint, länger) mit Acetylsalicylsäure/Clopidogrel behandelt werden muss und nach den heute vorliegenden Daten dann im Unfall- oder abdominalchirurgischen operationspflichtigen Notfall Sterblichkeiten bis in den zweistelligen Bereich durch Blutung oder Infarkt berichtet wurden.

28b.8 Fazit des Herzchirurgen

Bei Patienten mit koronarer Drei-Gefäß-Erkrankung ist die evidenzbasierte Therapie-Empfehlung an den Patienten eindeutig die möglichst komplette Revaskularisation mittels koronarer Bypassoperation mit mindestens arterieller Versorgung des RIVA. Auch und gerade nach ersten Kurzzeiterfahrungen mit der Mehr-Gefäß-Behandlung mit medikamentenbeschichteten Stents und den überraschenden mittelfristigen Ergebnissen sowie neu aufgetauchten Fragen ändert sich hieran nichts. Koronarchirurgie ist in grossen Serien in der hier geschilderten Konstellation mit proximalen Stenosen aller 3 Hauptäste der Intervention eindeutig in Prognose und Re-Interventionsrate überlegen, vermeidet Langzeitriskien wie aggressive Doppel-Antikoagulation bei älteren Patienten mit immanentem nicht-kardialen Operations- und Unfallrisiko und ist auch in fortgeschrittenem Lebensalter mit niedrigem kalkuliertem Risiko durchzuführen.

Literatur

- [1] Eagle KA et al., ACC/AHA 2004 guideline update for coronary artery bypass graft surgery: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1999 Guidelines for Coronary Artery Bypass Graft Surgery). *Circulation* (2004) 110(14), e340–437
- [2] Grover FL, Shroyer AL, Hammermeister KE. Calculating risk and outcome: the Veterans Affairs database. *Ann Thorac Surg* (1996) 62(5 Suppl), 6–11; discussion 31–32
- [3] Varnauskas E Twelve-year follow-up of survival in the randomized European Coronary Surgery Study. *N Engl J Med* (1988) 319, 332–337
- [4] Yusuf S et al., Effect of coronary artery bypass graft surgery on survival: overview of 10-year results from randomised trials by the Coronary Artery Bypass Graft Surgery Trialists Collaboration. *Lancet* (1994) 344 (8922), 563–570. Erratum in: *Lancet* (1994) 344 (8934), 1446
- [5] Myers WO et al., CASS Registry long term surgical survival. *Coronary Artery Surgery Study*. *J Am Coll Cardiol* (1999) 33, 488–498. Erratum in: *J Am Coll Cardiol* (1999) 34, 1363
- [6] Loop FD et al., Influence of the internal-mammary-artery graft on 10-year survival and other cardiac events. *N Engl J Med* (1986) 314, 1–6
- [7] Smith PK et al., Selection of surgical or percutaneous coronary intervention provides differential longevity benefit. *Ann Thorac Surg* (2006) 82, 1420–1428
- [8] Brett W, Guller U, Zerkowski HR, CABG versus PCI in Coronary Artery Disease: What is the evidence? *Cardiac surgery today* (2005) 2, 43–55
- [9] Hoffman SN et al., A meta-analysis of randomized controlled trials comparing coronary artery bypass graft with percutaneous transluminal coronary angioplasty: one-to eight-year outcomes. *J Am Coll Cardiol* (2003) 41, 1293–1304
- [10] Hueb W et al., The medicine, angioplasty, or surgery study (MASS-II): a randomized, controlled clinical trial of three therapeutic strategies for multivessel coronary artery disease: one-year results. *J Am Coll Cardiol* (2004) 43, 1743–1751

- [11] Adamina M et al., Propensity scores and the surgeon. *Br J Surg* (2006) 93, 389–394
- [12] Guller U, Surgical outcomes research based on administrative data: inferior or complementary to prospective randomized clinical trials? *World J Surg* (2006) 30, 255–266
- [13] Brener SJ et al., Propensity analysis of long-term survival after surgical or percutaneous revascularization in patients with multivessel coronary artery disease and high-risk features. *Circulation* (2004) 109, 2290–2295
- [14] Hannan EL et al., Long-term outcomes of coronary-artery bypass grafting versus stent implantation. *N Engl J Med* (2005) 352, 2174–2183
- [15] Babapulle MN et al., A hierarchical Bayesian meta-analysis of randomised clinical trials of drug-eluting stents. *Lancet* (2004) 364 (9434), 583–591.
- [16] Van Domburg RT et al., The impact of the introduction of drug-eluting stents on the clinical practice of surgical and percutaneous treatment of coronary artery disease. *Eur Heart J* (2005) 26, 675–681
- [17] Guyton RA, Coronary Artery Bypass is Superior to Drug-Eluting Stents in Multivessel Coronary Artery disease. *Ann Thorac Surg* (2006) 81, 1949–1957
- [18] Pfisterer M et al. (BASKET-LATE Investigators), Late clinical events after clopidogrel discontinuation may limit the benefit of drug-eluting stents: an observational study of drug-eluting versus bare-metal stents. *J Am Coll Cardiol* (2006) 48 (12), 2584–2591. Epub 2006 Nov 2.
- [19] Graham MM et al., Survival after coronary revascularization in the elderly. *Circulation* (2002) 105, 2378–2384
- [20] Matt P, Bernet F, Zerkowski HR, Herzchirurgie im fortgeschrittenen Lebensalter. *Dtsch Arztebl* (2005) 102, A 1056–1060

29 Evidenzbasierte Medizin am Beispiel des lumbalen Bandscheibenvorfalles

Josef Ramsbacher

29.1 Einleitende Übersicht

Die meisten Bandscheibenvorfälle sind degenerativ bedingt, wobei der Grad der Degeneration von der Leistungsfähigkeit des Gewebes und von der Größe der Belastung abhängig ist. Wegen der doppelten S-Form der Wirbelsäule und der damit verbundenen starken Belastung der Hals- und Lendenwirbelsäule, finden sich Bandscheibenvorfälle hauptsächlich in diesen beiden Abschnitten.

Die Bandscheibe besteht aus einem äußeren Faserring, dem Annulus fibrosus, der den weichen Kern, den Nucleus pulposus, einfasst. Der allmählich im Laufe des Lebens fortschreitende Flüssigkeitsverlust der gesamten Bandscheibe führt zu einer Strukturveränderung mit Elastizitätsverlust, und als Folge der „Materialermüdung“ kann der Faserring nachgiebig werden. Es kommt zu einer Bandscheibenvorwölbung (Protrusion). Wenn der Faserring reißt und es zum Austritt eines Teiles des Nucleus pulposus in den Wirbelkanal kommt, spricht man von einem Bandscheibenprolaps. Wird durch die Protrusion oder durch den Prolaps eine Nervenwurzel oder das Rückenmark komprimiert, so kommt es zu neurologischen Symptomen, deren Ausmaß von der Lage und Größe der Protrusion oder des Prolapses abhängig ist.

Die ersten Bandscheibenvorfälle wurde über eine Laminektomie bzw. über eine Hemilaminektomie entfernt. Die Langzeitergebnisse dieser Operationen waren jedoch – bedingt durch eine Reihe von Komplikationen, wie Infektionen, Rezidivbandscheibenvorfälle, Narbenbildung und Instabilitäten der Wirbelsäule – nicht signifikant besser als

bei konservativ behandelten Patienten [1]. Aus diesem Grund wurden im weiteren Verlauf neue Techniken entwickelt. Mittlerweile werden lumbale Bandscheibenvorfälle entweder mikrochirurgisch über eine interlamina-re Fensterung oder endoskopisch operiert. Neben diesen operativen Techniken wurde unter anderem auch die Chemonukleolyse entwickelt, bei der der Nucleus pulposus durch das über eine dünne Kanüle injizierte Chymopapain „verdaut“ wird. Durch diese Volumenreduktion kommt es zu einer Entlastung der neuralen Strukturen [2].

Bei der mikrochirurgischen Technik wird über einen ca. 3 cm langen Hautschnitt das Ligamentum flavum eröffnet und teilweise entfernt. Unter mikroskopischer Sicht werden anschließend der Bandscheibenvorfall und – falls nötig – ein Teil des Nucleus pulposus entfernt. Bei diesen Eingriffen ist nach wie vor die Narbenbildung im Bereich der Nervenwurzel und die Destruktion von Muskelgewebe ein Problem.

Bei den endoskopischen Techniken wird über einen kleinen lateralen Zugang ein Endoskop nach intraspinal vorgeschoben, wobei weder die Muskulatur noch das Ligamentum flavum verletzt werden. Der Nachteil dieser Technik besteht in ihrer eingeschränkten Anwendbarkeit, da nur ca. 25% aller lumbalen Bandscheibenvorfälle erfolgreich endoskopisch operiert werden können.

Trotz der verbesserten minimal-invasiven Techniken klagen in den USA von den ca. 200000 Patienten, die jährlich an einem Bandscheibenvorfall operiert werden, postoperativ zwischen 20 und 40% über dauerhafte neurologische Ausfälle oder Schmerzen.

29.2 Klinisches Szenario

Herr Siegfried, ein 45-jähriger Patient, bemerkte vor drei Wochen erstmalig eine leichte Fuß- und Großzehenheberschwäche sowie ein Taubheitsgefühl am Fußrücken und an der Großzehe rechts. Das klinische Bild wird durch starke Schmerzen im Bereich der Lendenwirbelsäule sowie im Bereich der Wade rechts komplettiert. Bei Belastung kommt es zusätzlich zu Kribbelmissempfindungen am Fußrücken und zu einer Zunahme der Schmerzen sowohl im Bereich der Lendenwirbelsäule als auch im Bereich der Wade rechts.

Im Lendenwirbelsäulenbereich bestehen beiderseits neben der Wirbelsäule Muskelverhärtungen und die Bewegung der Wirbelsäule ist schmerzbedingt eingeschränkt. Das Anheben des gestreckten rechten Beines ruft bei 35° einen Nervendehnungsschmerz hervor.

Anamnestisch ergibt sich, dass die Symptome vor drei Wochen bei starker Belastung plötzlich aufgetreten sind, und dass sie sich seitdem nicht verändert haben. Eine konservative Therapie erfolgte nur in Form von Analgetika (Paracetamol).

29.3 Kritische Analyse der vorhandenen Evidenz

Neben Expertenmeinungen, Kongressen und weiterführender Literatur ist heute das Internet, vor allem die MEDLINE-Datenbank, die wichtigste Informationsquelle.

In unserem Fall wurde, um die derzeit geltenden Standards der nötigen diagnostischen Verfahren zu finden, folgende Suche veranlasst:

- ▲ Suchbegriffe: „X-ray“ and „CT“ and „MRI“ and „Disc“ (Suchmodus: all fields). Ergebnis: 154 Veröffentlichungen.
- ▲ Weitere Eingrenzungen durch den Begriff: „Randomized“ – keine Dokumente vorhanden.

- ▲ Die Suche nach randomisierten kontrollierten Studien ergab: Es gibt in MEDLINE keine kontrollierten randomisierten Studien, die dem Evidenzgrad I oder II entsprechen würden. Dreiundvierzig veröffentlichte deutsch- und englischsprachige Reviews wurden gefunden, und in die nähere Auswahl gezogen.

Bei der Suche nach dem besten diagnostischen Vorgehen wurde zusammenfassend Folgendes gefunden:

- ▲ Alle bildgebenden Verfahren zeigen mit einer hohen Prävalenz abnorme Befunde in der Lendenwirbelsäule, die jedoch mehrheitlich keine Symptome hervorrufen und somit nur in der Verbindung mit den klinischen Daten zu einer zuverlässigen Aussage führen.

Aus diesem Grund empfiehlt die Dutch Society for Neurology die Untersuchung durch einen Spezialisten und weiterführende Diagnostik nur, wenn die Indikation zur Operation gestellt werden sollte oder wenn die Diagnose unsicher ist.

Folgenschwerer ist die Wahl der richtigen Behandlung. Hierbei lautet die Fragestellung: Welches Operationsverfahren – Intradiskale Technik, Mikrochirurgische Technik, Chemonukleolyse, Transforaminale Endoskopische Technik – bietet die größten Vorteile? Die Suche wird auf MEDLINE begrenzt. Die Suchterminologie sieht wie folgt aus:

- ▲ „Low back surgery“ and „randomized“. Ergebnis: 46 Veröffentlichungen (Suchmodus: all fields).
- ▲ Desweiteren ergaben die Suche in MEDLINE unter „clinical queries“ (im Suchmodus etiology, sensitivity) mit den Suchtermini: „Lumbar discectomy“ 241 Veröffentlichungen und mit dem Zusatz „randomized“ 23 Veröffentlichungen.

In dieser Liste der oben genannten Suchergebnisse wurde eine Veröffentlichung gefunden,

die besonders geeignet erscheint, die vorangehende Fragestellung zu beantworten. Der Abstract im Internet weist darauf hin, dass EbM angewandt worden ist und dass es sich um ein „Review“ handelt, das neun randomisierte Studien, sechs Meta-Analysen, Review-Artikel, eine Leitlinie der EbM (Vorgehen zu diesem Fall), 38 chirurgische Fallserien und 35 zusätzliche Referenzen beinhaltet. Veröffentlicht wurde der Artikel „Efficacy of lumbar discectomy and percutaneous treatment for lumbar disc herniation“ in „Sozial- und Präventivmedizin“ von Carl D. Stevens [3].

Bemerkenswert ist, dass eine chirurgische Indikation bei Lumbalgien ohne neurologische Ausfallerscheinung bis zum Caudaequina-Syndrom reicht, und dass die neurologische Untersuchung zur Identifizierung eines Bandscheibenvorfalles nur eine Sensitivität von 50% und eine Spezitivität von 70% hat. Bildgebende Verfahren werden sehr häufig durchgeführt, führen aber nur in einem niedrigen Prozentsatz zu einem operativen Vorgehen.

Bezüglich der Chemonukleolyse kommen die Autoren zu dem Schluss, dass Studien, die die chirurgische Discektomie mit der Chemonukleolyse vergleichen, einstimmig besagen, dass ein größerer Erfolg mit der gewöhnlichen „open“-Operationstechnik zu erzielen sei. Ausschlaggebend ist aber die Feststellung, dass die Rate der fehlerhaften Chemonukleolysen zwischen 20 und 56% liegt.

Über die perkutane Discektomie sind in der Literatur vorwiegend klinische Fallserien mit einer geringen Anzahl von Patienten zu finden. Es werden nur zwei randomisierte Studien genannt, wobei eine ein Spezialverfahren analysiert und die andere die perkutane Discektomie mit der Chemonukleolyse vergleicht.

Über Techniken wie Laser-Discektomie und endoskopische Discektomie existieren häufig nur vorläufige Studien. Die Effektivität dieser Techniken kann darum zurzeit mangels valider Daten nicht evaluiert werden.

29.4 Diskussion

Nachdem Langzeitergebnisse gezeigt haben, dass durch eine Laminektomie oder eine Hemilaminektomie in der Bandscheibenchirurgie keine signifikant besseren Ergebnisse erzielt werden als durch die konservative Therapie, wurde postuliert, dass die Eingriffe minimalisiert werden müssten [1; 4]. Hierbei konnte sich die mikrochirurgische Technik schnell durchsetzen. Im Rahmen der Umorientierung wurden aber auch viele alternative Verfahren, wie z.B. die Chemonukleolyse entwickelt. Diese perkutanen Operationsverfahren entsprachen dem Zeitgeist und waren sowohl bei den Medizinern als auch bei den Patienten sehr beliebt, da sie laut „Werbung“ schonender sind und dem Patienten eine schnellere Schmerzfreiheit garantieren. Nach den ersten Erfolgsmeldungen [5; 6], die auch in der Boulevardpresse erschienen, stieg der Druck auf die Chirurgen, diese Verfahren anzuwenden. Dadurch wurde schnell eine große Anzahl von Patienten mit diesen Techniken operiert, wobei sich aber allmählich herausstellte, dass diese Verfahren nur sehr bedingt erfolgreich einsetzbar sind [3; 7]. Trotzdem wurden diese Techniken mehr als zehn Jahre favorisiert, bevor die schlechten Ergebnisse zu einem Umdenken führte.

Die Gründe für die lange Vorherrschaft dieser Methoden liegen auf der Hand:

- ▲ Alle perkutanen Operationstechniken können im ambulanten Bereich eingesetzt werden.
- ▲ Trotz der großen Fallzahlen wurden nur wenige randomisierte Studien durchgeführt.
- ▲ Der Druck der Patienten, mit dieser Technik operiert zu werden, war sehr groß.
- ▲ Aus existenziellen Gründen sahen sich die Ärzte „gezwungen“, diese Techniken anzubieten.

29.5 Auflösung des klinischen Szenarios

Da es sich bei Herrn Siegfried um einen Bandscheibenvorfall handelt, bei dem der Annulus fibrosus bereits „gerissen“ ist, wird dem Patienten grundsätzlich von einem intradiskalen Vorgehen abgeraten, da bei all diesen Verfahren (Chemonukleolyse, perkutane endoskopische Laser Discektomie, etc.) ein intakter Annulus fibrosus Voraussetzung für die Indikation ist.

Eine transforaminale endoskopische Sequestrektomie erscheint ebenfalls wenig aussichtsreich, da erfahrungsgemäß, wie bei Herrn Siegfried vorhanden, eine Hypertrophie der Zwischenwirbelgelenke das Zwischenwirbelloch einengt und somit die Einführung des Endoskopes erschwert oder verhindert. Zudem erlaubt der enge Recessus lateralis keine sichere Sequestrektomie.

Als optimale Methode wird dem Patienten daher das mikrochirurgische Verfahren empfohlen. Mit dieser Methode ist eine komplette Entfernung des Sequesters ebenso wie die partielle Ausräumung des Zwischenwirbelraumes möglich. Das ist ausreichend, um den zuvor unter Druck stehenden Nerv zu entlasten. Der längere Krankenhausaufenthalt (1 Woche), die Gefahr der Narbenbildung sowie das Risiko einer seltenen Wirbelsäuleninstabilität sind Nachteile gegenüber den endoskopischen Verfahren. Diese werden aber durch die sichere und ausgiebige Beseitigung des sequestrierten Bandscheibenmaterials kompensiert.

Die Beantwortung der prozentualen Wahrscheinlichkeit eines guten postoperativen Ergebnisses ist trotzdem äußerst schwierig und sehr facettenreich. Abgesehen von den klinisch neurologischen Verschlechterungen ist die Zufriedenheit des Patienten der höchste Maßstab, wobei die Angaben in der Literatur zwischen 88 und 98% zufriedener Patienten schwanken.

Im klinischen Alltag bleibt dabei der psychosoziale Aspekt weitgehend unbeachtet, obwohl eine Reihe von Arbeiten auf die psychodynamischen Theorien der Schmerzentstehung, die Auswirkung von persönlichen Ereignissen sowie über

die multifaktorielle Ätiologie der Chronifizierung von Schmerzen hinweisen [8; 9]. Erst in den letzten Jahren wird häufiger, wie auch in unserer Klinik, ein interdisziplinäres Vorgehen bei der Behandlung von Patienten mit Bandscheibenvorfällen angeboten.

29.6 Schlussfolgerung

Die Neurochirurgie ist ein Fach mit multidisziplinärem Ansatz, das sich durch ständige Innovationen vonseiten der Technik kontinuierlich weiterentwickelt [10]. Wie uns schwer zu erkennen ist und auch oft in Fachzeitschriften erwähnt wird, fehlen häufig valide Studien, die eine Bewertung über die Effektivität zulassen [11].

In diesem Fach EbM anzuwenden ist problematisch, einerseits wegen der fortlaufenden Entwicklung, andererseits aufgrund der Schwierigkeit, valide Studien durchzuführen. Daher wird gefordert, durch Multicenterstudien mehr Klarheit über die Effektivität der einzelnen Operationsverfahren zu gewinnen. Nur dadurch kann der praktizierende Neurochirurg die einzelnen Operationstechniken vergleichen und den Patienten die optimale Therapie empfehlen.

Literatur

- [1] Weber H, Lumbar disc herniation. A controlled, prospective study with ten years of observation. *Spine*. (1983), 8, 131–140
- [2] Roggendorf W et al., Morphological alterations of the degenerated lumbar disc following chemonucleolysis with chymopain. *J Neurosurg*. (1984), 60, 518–522
- [3] Stevens CD et al., Efficacy of lumbar discectomy and percutaneous treatments for lumbar disc herniation. *Soz Präventivmed*. (1997), 42, 367–379
- [4] Ogon M et al., Minimal-invasive Zugänge und Operationsverfahren an der Lendenwirbelsäule. *Orthopäde*. (1997), 26, 553–561

- [5] Hoogland T, Percutaneous endoscopic discectomy. *J Neurosurg.* (1993), 79, 967–968
- [6] Yonezawa T et al., The system and procedures of percutaneous intradiscal laser nucleotomy. *Spine.* (1990), 15, 1175–1185
- [7] Dullerud R, Nakstad PH, Side effects and complications of automated percutaneous lumbar nucleotomy. *Neuroradiology.* (1997), 39, 282–285
- [8] Doering S, Sollner W, Zur Psychosomatik chronischer Schmerzen. *Orthopäde.* (1997), 26, 521–527
- [9] Heger S, Zur Psychosomatik des Failed-back-Syndroms: warum Rückenschmerzen chronifizieren. Plädoyer für einen zeitgemässen Umgang mit den Lumbago-Ischi-algie-Syndromen. *Nervenarzt.* (1999), 70, 225–232
- [10] Powell M, Recent advances: neurosurgery. *BMJ.* (1999), 318, 35–38
- [11] Wilkins RH, Science in neurosurgery: the importance of the scientific method. *Neurosurgery.* (1998), 42, 687–691

30 Schlaganfall und Rehabilitation

Günter Seidel

30.1 Einleitung

Der Schlaganfall ist die dritthäufigste Todesursache in Deutschland. Die Zahl der jährlichen Neuerkrankungen liegt in Zentraleuropa bei ca. 150 pro 100 000 Einwohner. Altersabhängig nimmt die Häufigkeit deutlich zu und übersteigt im höheren Alter die Ereignisrate des akuten Myokardinfarktes [1]. Die Letalität wird im ersten Monat nach dem Schlaganfall mit ca. 19% angegeben und steigt auf 37% im ersten Jahr [2]. Die überlebenden Patienten haben eine mittlere Lebenserwartung von 7,3 Jahren und bleiben in 25% der Fälle auf den Rollstuhl angewiesen oder sind bettlägerig. Motorische Defizite sind mit Abstand die häufigsten Symptome und kommen bei etwa 90% der Schlaganfallpatienten vor. Die stärkste Rückbildung kann in den ersten 12 Wochen erwartet werden, bei schweren Ausfällen erstreckt sich die Rückbildungsphase über mehrere Monate bis vereinzelt auch über Jahre. Die motorische Rehabilitation spielt somit eine wichtige Rolle zur Wiedererlangung der Unabhängigkeit des Patienten.

Neurobiologische Grundlage der motorischen Rehabilitation ist die neuronale Plastizität, mit der die Fähigkeit des Hirns beschrieben wird, sich durch Umorganisation und an veränderte Umgebungsbedingungen anzupassen [3]. In den letzten Jahren wurde eine große Anzahl von kontrollierten Studien publiziert, die zeigten, dass neue Behandlungstechniken in der Lage sind, die Plastizität des Hirns positiv zu beeinflussen.

30.2 Klinisches Szenario

Der 68-jährige Herr Müller hat einen subkortikalen Teilinfarkt im Versorgungsgebiet der A. cerebri media mit brachiofazial-betonter Hemiparese links auf dem Boden einer kardialen Hirnembolie durch Vorhofflimmern erlitten. Nach Akuttherapie und Einleitung der Sekundärprävention auf der Schlaganfall-Spezialstation (Stroke Unit) befindet er sich nun 5 Tage nach Infarkteintritt in einem stabilen klinischen Zustand. Die Hirnschädigung hat zu einer kompletten Armparese mit aufgehobener Greiffunktion und einer Beinparese mit erhaltener Steh-, aber aufgehobener Gehfähigkeit geführt. Die Sensibilität ist weitgehend intakt geblieben und neuropsychologische Defizite finden sich nicht. Der Patient möchte die optimale rehabilitative Behandlung für sich in Anspruch nehmen und fragt nach einer eingehenden Beratung diesbezüglich.

30.3 Evidenz physiotherapeutischer rehabilitativer Therapieprinzipien

Wesentliche Effekte auf das Wiedererlangen motorischer Funktionen hat die Intensität des Trainings, die Häufigkeit der Übungswiederholungen und die Umsetzung der Trainingsleistungen in Alltagsaktivitäten, wobei der Aufgabenspezifität eine entscheidende Bedeutung zuzukommen scheint [4; 5]. Den neuen physiotherapeutischen Verfahren [6; 7] stehen die klassischen Verfahren wie die Techniken nach Bobath und Votja sowie die propriozeptive neuromuskuläre Fazilitierung (PNF) gegenüber. Die motorische Rehabilita-

tion kann durch verschiedenen Pharmaka positiv beeinflusst werden. Im Rahmen der motorischen Rehabilitation werden pragmatisch die Rehabilitation der oberen und die der unteren Extremität unterschieden. Im Folgenden werden verschiedene rehabilitative Verfahren dargestellt:

Bei der Rehabilitation der Arm- und Handfunktion (s. Tab. 30.1) in der Akut- und Subakutphase hat das **repetitive sensomotorische Training**, bei dem beispielsweise die Schulter oder die Hand repetitiv durch eine Mechanik bewegt wird [8], eine signifikante Verbesserung verschiedener funktioneller Skalen der Hand- und Armfunktion insbesondere bei schwer betroffenen Patienten mit Neglekt erbracht, wobei diese Verbesserungen auch nach fünf Jahren noch nachweisbar waren [9]. Ob die Verbesserung der Skalen auch einen Effekt auf das Alltagsleben der Betroffenen hatte, wurde nicht berichtet.

Die **Robot-assistierte Rehabilitation** (MIT-Manus) verbessert insbesondere die Funktion der proximalen Armmuskulatur. Verschiedene Arbeitsgruppen konnten nachweisen, dass diese Verfahren sowohl im Vergleich zur Placebo-Behandlung [10; 11] als auch im Vergleich zur Bobath-Behandlung [12] wirksamer sind und dass Patienten sowohl im akuten als auch im chronischen Stadium profitieren.

Die **therapeutische elektrische Stimulation** nutzt elektrischen Strom, um Muskeln direkt zur Kontraktion anzuregen. Dieses Verfahren kann EMG-getriggert durchgeführt werden (Biofeedback), wobei geringe Willküraktivität in den beübten paretischen Muskeln zu einer Elektrostimulation mit großem Bewegungseffekt führt. Am häufigsten wurde die Unterarmmuskulatur stimuliert, um die Handgelenksexension zu verbessern. In zwei Meta-Analysen, die bis zu sechs randomisierte klinische Studien nutzten, konnte der positive Effekt auf die Handgelenksexension belegt werden [13; 14], wobei aber die Alltagsrelevanz der Verbesserung nicht

gezeigt werden konnte. Neben der Verbesserung der Beweglichkeit sind positive Effekte auf die Spastik beschrieben worden. Bei Applikation der elektrischen Stimulation an den Mm. deltoideus und supraspinatus ist darüber hinaus über eine Abnahme der Subluxation des Schultergelenks berichtet worden. Dies konnte in einer Meta-Analyse signifikant belegt werden [7]. In einem systemischen Review zum reinem **EMG-Biofeedback** (ohne therapeutische elektrische Stimulation) konnte kein Zugewinn an selektiver Willkürinnervation und Armfunktion gezeigt werden [15].

Für das **Armfähigkeitstraining**, bei dem Patienten mit leichten Armparesen spezielle sensomotorische Fähigkeiten üben, konnte ein nachhaltig positiver Effekt auf alltagsähnliche Tätigkeiten nachgewiesen werden [16].

Das wiederholte Vorstellen von Bewegungsabläufen (**mentales Training**) verbessert sowohl Planung als auch Ausführung zuvor trainierter als auch neuer Aufgaben [17], wobei ein effektives Arbeitsgedächtnis Voraussetzung für die Behandlung ist. Die Evidenzbasis dieser Therapieform ist bisher gering, doch scheint hier ein vielversprechendes Verfahren vorzuliegen.

Auch das alleinige bewusste Sehen von Bewegungsabläufen hat einen positiven Effekt auf die motorische Rehabilitation [18], wobei dieser neue Ansatz zurzeit intensiv beforscht wird und der therapeutische Wert für die Praxis der motorischen Rehabilitation noch nicht abzuschätzen ist.

Das Prinzip des „**forcierten Gebrauchs**“ (Constraint Induced (CI) Movement Therapy, sog. Taub-Training) nutzt einen lerntheoretischen Ansatz im chronischen Stadium nach Schlaganfall zur motorischen Rehabilitation [19; 20]. Bei diesem Verfahren wird der gesunde Arm immobilisiert und der paretische Arm nicht nur während der Therapiestunden, sondern auch bei alltäglichen Verrichtungen vermehrt eingesetzt. Hierdurch

soll die „verlernte“ Benutzung des betroffenen Arms wiedererlernt werden. Die Wirksamkeit dieses Verfahrens im chronischen Stadium nach Schlaganfall konnte in kontrollierten Studien belegt werden [21; 22] und war auch zwei Jahre nach dem Training noch nachweisbar. Der zeitliche Aufwand und die Anforderungen an den Patienten sind hoch, wobei ein tägliches Trainingspensum von sechs Stunden pro Tag besser als 3 Stunden pro Tag zu sein scheint [23]. Die zeitlich verlängerte, aber weniger intensive Behandlung mit einer Therapiephase von 10 Wochen (dreimal pro Woche strukturierte Therapie mit Immobilisation der gesunden Extremität für fünf Stunden pro Tag an fünf Tagen in der Woche) führt zu besseren motorischen Scores als die durchgehende intensive Behandlung [24]. Die überlegene Effektivität des Taub-Trainings im Vergleich zur gleich intensiven Bobath-Therapie konnte belegt werden [25]. Der Einsatz des Verfahrens im akuten Stadium der Hirnläsion nach Schlaganfall wird nicht empfohlen, weil es tierexperimentelle Hinweis für eine Vergrößerung der Hirnläsion gibt [26].

Auch bei der motorischen Rehabilitation der Beinfunktion und des Laufens sollten aufgabenspezifische repetitive Übungsbehandlungen die Therapie dominieren. Vor diesem Hintergrund hat die Einführung der Laufbandtherapie mit partieller Körpergewichtsentlastung einen erheblichen Fortschritt erbracht, weil auch noch rollstuhlbedürftige Patienten gurtgesichert und entsprechend dem Paresegrad entlastet an dieser **Lokomotionstherapie** teilnehmen können. Bei Patienten, die bereits wieder selbstständige Gehfähigkeit erlangt haben, aber noch unsicher sind, hilft die gurtgesicherte Lokomotionstherapie, Ganggeschwindigkeit und Ausdauer zu erhöhen [27; 28]. Förderlich für die weitere Gangrehabilitation ist der **Einsatz von Hilfsmitteln** wie Stock oder Rollator und der Gebrauch von Orthesen zur Stabilisierung der Sprunggelenks- und Kniege-

lenksstabilität. Diese Hilfsmittel haben ihren positiven Einfluss auf die Gangrehabilitation durch die Erhöhung der Gangsicherheit. Alternativ zu orthetischen Versorgung des Spitzfußes bietet sich die **funktionelle Elektrostimulation** des M. tibialis oder des N. peroneus an. Die funktionelle Elektrostimulation verbesserte in einer randomisierten Studie auch die Spastik sowie die Gehfähigkeit und erhöhte den Anteil der Patienten, die nach der Therapie zu Hause leben konnten, im Vergleich zur Standard-Krankengymnastik [29].

30.4 Klinisches Szenario

Herr Müller fragt, ob es neben den erwähnten physikalischen Therapieformen auch medikamentöse Möglichkeiten gebe, die motorische Rehabilitation zu verbessern.

30.5 Evidenz pharmakologischer rehabilitativer Therapieprinzipien

Eine Beeinflussung der Neurotransmitterkonzentrationen und die Stimulation verschiedener Rezeptoren im Hirn können die Geschwindigkeit und das Ausmaß der motorischen Rehabilitation positiv, aber auch negativ beeinflussen.

In früheren Studien zur motorischen Rehabilitation von Schlaganfallpatienten hat **D-Amphetamin** einen positiven Effekt gezeigt, wenn die Einnahme unmittelbar vor intensiver repetitiver Physiotherapie erfolgte [30; 31]. Aktuellere Arbeiten konnten den positiven Effekt nicht immer nachweisen [32]. Negative Effekte der Behandlung können tachykarde Herzrhythmusstörungen und hypertensive Krisen sein, was insbesondere bei Schlaganfallpatienten mit kardialen Vorerkrankungen und kardiovaskulärem Risikoprofil limitierend bei der Behandlung ist.

Entsprechend müssen diese Parameter engmaschig während der Therapie kontrolliert werden. Unter kontrollierten Therapiebedingungen ist eine Suchtentwicklung nicht zu befürchten. Als Präparat bietet sich z.B. Methylphenidat (Ritalin) 20–40 mg/d an. Der Therapieeffekt setzt bereits nach 1–2 Tagen ein. Die Therapiedauer hängt vom Ansprechen und der Verträglichkeit ab. Optimales Zeitintervall sind die ersten 8–12 Wochen nach Schlaganfalleintritt, eine Behandlung länger als 3 Monate wird nicht empfohlen.

Auf das zentrale Belohnungssystem wirkt u.a. **L-Dopa**. Dies Substanz, die für die Behandlung von extrapyramidalen Erkrankungen zugelassen ist, konnte auch in einer placebokontrollierten Doppelblindstudie in einer Dosierung von 100 mg/d eine Verbesserung der Gehfähigkeit und der Armmotorik bewirken [33]. Günstig im Vergleich zum D-Amphetamin sind die deutlich geringer ausgeprägten kardiovaskulären Nebenwirkungen. Ein Vergleich der Wirksamkeit zwischen D-Amphetamin und L-Dopa wurde bisher nicht durchgeführt.

Die Depression ist häufig nach Schlaganfall zu finden (bis zu 40%) und kann reaktiv durch das Bewusstwerden der Behinderung oder durch strategische Läsionsmuster im Hirn verursacht sein [34]. **Antidepressiva** verändern die Neurotransmitterspiegel im Hirn. Für den Serotoninwiederaufnahmehemmer (SSRI) Fluoxetin in einer Dosierung von 20 mg/d konnte in einer placebokontrollierten Studie gezeigt werden, dass unabhängig von seinem antidepressiven Effekt eine Verbesserung der Gehfähigkeit und der Selbstständigkeit bei alltäglichen Verrichtungen erreicht werden konnte [35]. Zu beachten ist, dass Fluoxetin, wie auch andere SSRI, über das Cytochrom P450 Isoenzym in der Leber metabolisiert wird, was zu Medikamenten-Interaktionen führen kann. Dieser Zusammenhang kann insbesondere bei der Komedikation mit sekundärpräventiven Medikamenten wie den Thienopyridinen oder dem Phenprocoumon klinisch relevant werden.

Tab. 30.1: Therapieverfahren in der Armrehabilitation (modifiziert [36])

Verfahren	Indikation	Positiver Therapieeffekt	Evidenzstärke
Repetitive sensorimotorische Training	Mittelschwere bis schwere Parese, reduzierte selektive Innervation	Selektive Beweglichkeit (z.T. auch Funktion) in subakuter Phase	1b, Ü
Robot-assistierte Rehabilitation	Mittelschwere bis schwere Parese, reduzierte selektive Innervation	Selektive Beweglichkeit (z.T. auch Funktion) in akuter und chronischer Phase	1b, Ü
Therapeutische elektrische Stimulation ± EMG-Biofeedback	Mittelschwere bis schwere Parese, reduzierte selektive Innervation	Selektive Beweglichkeit (z.T. auch Funktion) in akuter und chronischer Phase	1a, Ü
Armfähigkeitstraining	Leichte Parese, reduzierte sensorimotorische Effizienz	Geschwindigkeit und Präzision der Armmotorik in subakuter und früher chronischer Phase	1b
Taub-Training	Gelernter Nichtgebrauch/partielle funktionelle Defizite	Einsatz/Funktion des Armes im Alltag in subakuter und chronischer Phase	1b, Ü

1a Evidenz aufgrund von Meta-Analysen randomisierter kontrollierter Studien (RKS), **1b** Evidenz aufgrund ≥ 1 RKS, **Ü** ≥ 1 RKS mit therapeutischer Überlegenheit gegenüber einer zeitlich gleich intensiven Kontrolltherapie.

30.6 Klinisches Szenario

Nach der ausführlichen Aufklärung über die rehabilitativen Therapieformen wird Herrn Müller eine multidisziplinäre Therapie (Physio- und Ergotherapie), bestehend aus einem L-Dopa-unterstützten repetitiven sensomotorischen Training der oberen Extremität und einem Laufbandtraining für die Beinmotorik – kombiniert mit der Bobath Therapie – vorgeschlagen.

Literatur

- [1] Rothwell PM et al., Population-based study of event-rate, incidence, case fatality, and mortality for all acute vascular events in all arterial territories (Oxford Vascular Study). *Lancet*. (2005), 366, 1773–1783
- [2] Kolominsky-Rabas PL et al., A prospective community-based study of stroke in Germany – the Erlangen Stroke Project (ESPro): incidence and case fatality at 1, 3, and 12 months. *Stroke*. (1998), 29, 2501–2506
- [3] Liepert J, Hamzei F, Weiller C, Lesion-induced and training-induced brain reorganization. *Restor Neurol Neurosci*. (2004), 22, 269–277
- [4] Langhammer B, Stanghelle JK, Bobath or motor relearning programme? A comparison of two different approaches of physiotherapy in stroke rehabilitation: a randomized controlled study. *Clin Rehabil*. (2000), 14, 361–369
- [5] Kwakkel G et al., Effects of augmented exercise therapy time after stroke: a meta-analysis. *Stroke*. (2004), 35, 2529–2539
- [6] Nelles G (2005) Motorische Rehabilitation nach Schlaganfall. Leitlinien der Deutschen Gesellschaft für Neurologie. Thieme, Stuttgart
- [7] Barreca S et al., Treatment interventions for the paretic upper limb of stroke survivors: a critical review. *Neurorehabil Neural Repair*. (2003), 17, 220–226
- [8] Feys HM et al., Effect of a therapeutic intervention for the hemiplegic upper limb in the acute phase after stroke: a single-blind, randomized, controlled multicenter trial. *Stroke*. (1998), 29, 785–792
- [9] Feys H et al., Early and repetitive stimulation of the arm can substantially improve the long-term outcome after stroke: a 5-year follow-up study of a randomized trial. *Stroke*. (2004), 35, 924–929
- [10] Volpe BT et al., A novel approach to stroke rehabilitation: robot-aided sensorimotor stimulation. *Neurology*. (2000), 54, 1938–1944
- [11] Fasoli SE et al., Robotic therapy for chronic motor impairments after stroke: Follow-up results. *Arch Phys Med Rehabil*. (2004), 85, 1106–1111
- [12] Lum PS et al., Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Arch Phys Med Rehabil*. (2002), 83, 952–959
- [13] de Kroon Jr et al., Therapeutic electrical stimulation to improve motor control and functional abilities of the upper extremity after stroke: a systematic review. *Clin Rehabil*. (2002), 16, 350–360
- [14] Bolton DA, Cauraugh JH, Hausenblas HA, Electromyogram-triggered neuromuscular stimulation and stroke motor recovery of arm/hand functions: a meta-analysis. *J Neurol Sci*. (2004), 223, 121–127
- [15] Moreland J, Thomson MA, Efficacy of electromyographic biofeedback compared with conventional physical therapy for upper-extremity function in patients following stroke: a research overview and meta-analysis. *Phys Ther*. (1994), 74, 534–543
- [16] Platz T et al., Arm ability training for stroke and traumatic brain injury patients with mild arm paresis: a single-blind, randomized, controlled trial. *Arch Phys Med Rehabil*. (2001), 82, 961–968
- [17] Liu KP et al., Mental imagery for promoting relearning for people after stroke: a randomized controlled trial. *Arch Phys Med Rehabil*. (2004), 85, 1403–1408
- [18] Buccino G, Solodkin A, Small SL, Functions of the mirror neuron system: implications for neurorehabilitation. *Cogn Behav Neurol*. (2006), 19, 55–63
- [19] Wolf SL et al., Forced use of hemiplegic upper extremities to reverse the effect of learned nonuse among chronic stroke and head-injured patients. *Exp Neurol*. (1989), 104, 125–132
- [20] Taub E et al., Technique to improve chronic motor deficit after stroke. *Arch Phys Med Rehabil*. (1993), 74, 347–354

- [21] Taub E, Crago JE, Uswatte G, Constraint-induced movement therapy: a new approach to treatment in physical rehabilitation. *Rehabil Psychol.* (1998), 43, 152–170
- [22] Wittenberg GF et al., Constraint-induced therapy in stroke: magnetic-stimulation motor maps and cerebral activation. *Neurorehabil Neural Repair.* (2003), 17, 48–57
- [23] Sterr A et al., Longer versus shorter daily constraint-induced movement therapy of chronic hemiparesis: an exploratory study. *Arch Phys Med Rehabil.* (2002), 83, 1374–1377
- [24] Page SJ et al., Efficacy of modified constraint-induced movement therapy in chronic stroke: a single-blinded randomized controlled trial. *Arch Phys Med Rehabil.* (2004), 85, 14–18
- [25] van der Lee JH et al., Forced use of the upper extremity in chronic stroke patients: results from a single-blind randomized clinical trial. *Stroke.* (1999), 30, 2369–2375
- [26] DeBow SB et al., Immediate constraint-induced movement therapy causes local hyperthermia that exacerbates cerebral cortical injury in rats. *Can J Physiol Pharmacol.* (2004), 82, 231–237
- [27] Pohl M et al., Speed-dependent treadmill training in ambulatory hemiparetic stroke patients: a randomized controlled trial. *Stroke.* (2002), 33, 553–558
- [28] Eich HJ et al., Aerobic treadmill plus Bobath walking training improves walking in subacute stroke: a randomized controlled trial. *Clin Rehabil.* (2004), 18, 640–651
- [29] Yan T, Hui-Chan CW, Li LS, Functional electrical stimulation improves motor recovery of the lower extremity and walking ability of subjects with first acute stroke: a randomized placebo-controlled trial. *Stroke.* (2005), 36, 80–85
- [30] Walker-Batson D et al., Amphetamine paired with physical therapy accelerates motor recovery after stroke. Further evidence. *Stroke.* (1995), 26, 2254–2259
- [31] Crisostomo EA et al., Evidence that amphetamine with physical therapy promotes recovery of motor function in stroke patients. *Ann Neurol.* (1988), 23, 94–97
- [32] Treig T et al., No benefit from D-amphetamine when added to physiotherapy after stroke: a randomized, placebo-controlled study. *Clin Rehabil.* (2003), 17, 590–599
- [33] Scheidtmann K et al., Effect of levodopa in combination with physiotherapy on functional motor recovery after stroke: a prospective, randomised, double-blind study. *Lancet.* (2001), 358, 787–790
- [34] Verdelho A et al., Depressive symptoms after stroke and relationship with dementia: A three-year follow-up study. *Neurology.* (2004), 62, 905–911
- [35] Dam M et al., Effects of fluoxetine and maprotiline on functional recovery in poststroke hemiplegic patients undergoing rehabilitation therapy. *Stroke.* (1996), 27, 1211–1214
- [36] Platz T, Evidenzbasierte Armrehabilitation. Eine systematische Literaturübersicht. *Nervenarzt.* (2003), 74, 841–849

31 Evidenzbasierte Medizin in der Pädiatrie – Beispiel Pseudokrupp

Johannes Forster, Dirk Bassler

Kinder sind keine kleinen Erwachsenen. Ihre Anatomie, Physiologie, Immunologie und pharmakokinetischen Daten sind anders als bei Erwachsenen. Sie machen daher Erkrankungen durch, die im Erwachsenenalter in gleicher Form nicht vorkommen.

Eine solche Erkrankung, die häufig vorkommt, Eltern in höchste Bedrängnis bringt und oft zur Krankenhausaufnahme führt, ist die Laryngotracheitis.

Der so genannte Pseudokrupp (engl. pseudocroup; auch einfach Krupp genannt, da der diphtherische Krupp praktisch nicht mehr vorkommt und zahlenmäßig keine ernste Differenzialdiagnose darstellt) war in den 1970er- und frühen 1980er-Jahren ein wichtiges öffentliches Thema. Damals glaubte man, dass die winterliche Luftverschmutzung Pseudokrupp auslöse. Lokal vorgefundene Assoziationen hoher Stickoxyd- (NO_x) und Staubbelastung mit hoher Inzidenz der Erkrankung schienen dies zu beweisen. Mittlerweile ist durch sorgfältige Analyse dieser Daten und infektionsepidemiologischer Befunde klar, dass in der Mehrzahl der Fälle Virusinfektionen der Luftwege ursächlich sind. Luftverschmutzung und familiäre Belastung sind dagegen nur Kofaktoren, die die laryngotracheale Hyperreagibilität verstärken.

31.1 Das Versorgungsproblem

Für viele Familien beginnt die Erkrankung „aus heiterem Himmel“, vorwiegend in der Nacht, mit bellendem Husten und inspiratorischem Stridor, der bei mittelschweren Fällen auch in Ruhe fortbesteht. Die Patienten

werden meist in diesem Stadium dem niedergelassenen Arzt vorgestellt. Dieser versucht nun zu verhindern, dass die laryngotracheale Schleimhautschwellung fortbesteht und eine Krankenhausaufnahme wegen drohender Hypoxie oder (letztes Stadium) Ateminsuffizienz erforderlich wird.

Die Behandlung dieses akuten Krankheitsbildes folgt seit sechs Jahren [1] den gleichen Prinzipien: es wird systemisch ein Corticoid verabreicht. Diese Behandlung ist für die leichten und mittelschweren Fälle ausreichend. Bei akut schweren Fällen erhält das Kind eine Suprarenin-Inhalation, diese wird auch im Verlauf wiederholt eingesetzt, wenn es zu rekurrenden Symptomen kommt. Nur in seltenen Fällen ist eine intensiv-medizinische Behandlung einschließlich der endotrachealen Intubation notwendig. Die Erkrankung dauert üblicherweise 1 bis 2 Tage. Die Symptome sind nachts, vor allem in der Nacht nach Erkrankungsbeginn, besonders ausgeprägt.

Krupp-Ereignisse wiederholen sich sowohl bei einzelnen Kindern wie auch familiär. Kinderkrankenhäuser und Kinderärzte geben daher Merkblätter an die betroffenen Familien, in denen das Vorgehen beim leichten Krupp-Anfall beschrieben ist: Den Eltern wird geraten, das Kind entweder feuchter Luft (Wasserdampf) oder kalter Luft (Platznehmen am geöffneten Fenster; die Erkrankungen treten ja im Winter bei den entsprechend niedrigen Außenlufttemperaturen auf) auszusetzen. Außerdem sollen die Eltern ein Kortison-Präparat bereithalten, welches bei anhaltendem Symptom dem Kind zu geben ist (AWMF-Leitlinien).

31.2 Klinisches Problem

Im März des Jahres 2006 erregt ein Kinderarzt in seinem Qualitätszirkel Aufsehen, der behauptet, dass die Anwendung feuchter Luft bei Krupp-Patienten „nichts bringt“. Man müsse sogar befürchten, dass die Empfehlung dieser nutzlosen Maßnahme zur Verschleppung wirksamer Behandlung führe und damit potenziell zur Gefährdung der Patienten mit mittelschwerem und schwerem Krupp. Dies sei in einer randomisierten Studie festgestellt worden, sodass das entsprechende Merkblatt sicherheitshalber unverzüglich geändert werden müsse.

31.3 Festlegung der Fragestellung

Den Teilnehmern des Qualitätszirkels erscheint es durchaus notwendig, diesem Befund nachzugehen und ihn gegebenenfalls auch im Merkblatt zu berücksichtigen. Andererseits sind ab März nicht mehr allzu viele Krupp-Fälle zu erwarten, sodass ein halbes Jahr für die Klärung der Fragestellung zur Verfügung steht.

Die Gruppe will zur Lösung des Problems hauptsächlich folgende Fragen bearbeiten:

- ▲ Führt die Anwendung von feuchter und/oder kalter Atemluft bei Kindern mit leichtem und mittelschwerem Krupp zur Besserung der Atemnot?
- ▲ Gibt es Hinweise, dass die Atemluft-Therapie tatsächlich zu einer gefährlichen Verschleppung nützlicher Maßnahmen geführt hat?

31.4 Die Evidenz suchen und bewerten

31.4.1 Wirksamkeit der Atemluft-Anfeuchtung

Zur ersten Frage finden sich zwei randomisierte Studien: die neueste, die für das initiale Aufsehen gesorgt hatte [2], und eine weite-

re [3], welche in PubMed aufgrund der Suchworte „Croup“ und „Mist“ gefunden wurde. Eine weitere Studie mit sehr geringer Patientenzahl war nur in Abstract-Form erhältlich [4].

Die beiden erstgenannten Studien wurden von je einem Mitglied des Qualitätszirkels dargestellt und von den Mitgliedern gemeinsam bewertet.

Die Gruppe hält die Studien, obwohl sie in Kanada durchgeführt wurden, für relevant für die eigene Praxis. Insbesondere die kurze Zeit vom Symptombeginn bis zur Vorstellung spricht dafür, dass die Notfallambulanzen der beteiligten Krankenhäuser als primäre Anlaufstelle für die Krupp-Patienten dienen (Primär-Versorgung).

Der Nutzeffekt einer Anfeuchtung der eingeatmeten Luft wird anhand der Studien als niedrig eingeschätzt. In der Neto-Studie ist die Besserung ganz offensichtlich der Dexamethason-Behandlung zuzuschreiben; und dieser Effekt ist so groß, dass daneben kein zusätzlicher Effekt durch die Luftanfeuchtung eintritt. Wenn – und das gilt für beide Studien – bei einer Erhöhung der Patientenzahl ein Therapieeffekt nachweisbar wäre, so wäre er doch von sehr geringem Ausmaß.

Obwohl dies nicht Gegenstand der beiden Studien war, ergibt sich aus der großen Zahl der vor-untersuchten Kinder und der darin enthaltenen ebenfalls großen Zahl von Kindern mit einem nur noch sehr leichten Krupp bei Vorstellung in der Notfallambulanz der Hinweis, dass auch die (auf dem Weg zum Krankenhaus verbrachte) Zeit einen heilsamen Effekt haben könnte. Dieses entspricht auch der Praxiserfahrung, dass sich die Erst-Symptome offensichtlich bei sehr vielen Patienten bis zur Vorstellung gebessert haben (unmittelbare Evidenz aufgrund von Eltern-Aussagen).

Zusammenfassend wird also die Wirkstärke des Kortison-Präparates, die ja im Vergleich zu Placebo nachgewiesen ist, eindeu-

Tab 31.1: Darstellung und Bewertung der beiden aktuellen kontrollierten Studien

	Scolnik et al., 2006 [2]	Neto et al., 2002 [3]
Design	randomisierte, einfach verblindete Studie 3 x 46 Kinder	randomisierte, einfach verblindete Studie 2 x 35 Kinder
Patienten	Notaufnahme Uni-Klinik Toronto 2 Jahre (3 Monate – 10 Jahre) Krupp-Score* Mittel: 3 Punkte Symptom bis Behandlungsbeginn: Mittel: 25 Min.	Notaufnahme Eastern Ontario Hospital 1¾ Jahre (3 Monate – 6 Jahre) Krupp-Score* Mittel: 4 Punkte
Behandlung	1. Vernebler vorhalten (keine Änderung der Luftfeuchte) 2. O ₂ -Maske und definitiv 40% Luftfeuchte 3. O ₂ -Maske und definitiv 100% Luftfeuchte	Alle Patienten: Dexamethason oral 1. „misty stick“ 2. kein „misty stick“
Ergebnis	Keine Unterschiede hinsichtlich Krankenhausaufnahmen (< 5%) Dexamethason (85%) Epinephrin (8%)	Keine Unterschiede über 2 Stunden in Croup Score, Sauerstoff-Sättigung, Puls- und Atemfrequenz, Eltern-Urteil über Zustand des Kindes
Wichtiges	A) Von 987 voruntersuchten Kindern hatten bis zum Studienbeginn 438 einen Croup Score < 2 B) Besserung des Croup Scores binnen 60 Min. in allen Gruppen um durchschnittlich 1 Punkt	A) Von 648 voruntersuchten Kindern hatten 473 einen Croup Score < 2. B) Die Kinder hatten nach 2 Stunden eine Besserung des Croup Scores von durchschnittlich 2–7 Punkten.

* Westley-Score: 0–1 Punkte: leichter, 2–6 mittelschwerer, 7–17 schwerer Croup

tig höher eingestuft als die Exposition zu befeuchteter oder kühler Luft.

31.4.2 Gefährdung von Kindern durch Verzögerung der Kortison-Gabe aufgrund der primären Anwendung von feuchter Luft

Die Gruppe, die diese Fragestellung erarbeiten wollte, startete zunächst mit einer Suchstrategie, die auf der Wortkombination „Croup“, „Hospitalization“ und „Mist“ OR „Humidi“* basierte, erhielt aber keine auch nur ansatzweise aussagekräftige Studie. Bei der weiteren Suche mit „Croup“ als MeSH-Term, wurden zwei relevante Publikationen gefunden: Zum einen die Arbeit von Segal et al. „Croup Hospitalization in Ontario“ [5], in der ein klarer Rückgang der Hospitalisie-

rungszahlen ab dem Winter 93/94 zu verzeichnen ist – dieser Winter folgte der Bekanntmachung, dass systemische Dexamethason-Behandlung einen eindeutig die Hospitalisierungsrate vermindernenden Effekt bei Krupp-Erkrankungen hat. Zum anderen fand sich eine Fallserie von Greally et al. [6], in der von an Krupp erkrankten Kindern berichtet wird, die mit Verbrühungen dem Arzt vorgestellt wurden.

Die retrospektive Studie von Segal et al. [5] wird von der Gruppe so gedeutet, dass die systemische Kortison-Behandlung den Grundstein einer erfolgreicher Krupp-Behandlung darstellt. Klinik-Ärzte könnten nun von ihrem Krankengut her schließen, dass Luftanfeuchtung nicht nur nichts bewirkt (ihre Patienten sind ja in einem schlechten Zustand, obwohl sie Luftanfeuchtung erhalten hatten), sondern auch durch

die Verzögerung wirksamer Therapie die Kinder schädigen kann. Für diese Kollegen zeigt der Befund, dass auch während des einstündigen Zuwartens vor der Dexamethason-Gabe in der Scolnik-Studie die Croup Scores gesunken waren – somit ist ganz klar, dass die bisherige Empfehlung keine klinischen Nachteile für die Patienten hatte. Für Klinik-Ärzte bedeutet dies: Diejenigen Patienten, denen es nach einer Erstberatung schon besser geht, kommen nicht zu ihnen. Diejenigen, die kommen, tun dies wegen des individuell schlechten Krankheitsverlaufes und nicht, weil irgendeine der Erstempfehlungen „schädlich“ gewesen wäre (Attribution-Bias).

Die Fallserie von Greally et al. [6] belegt allerdings, dass direkte Risiken bestehen, z.B. bei der direkten Inhalation über erhitztem und kochendem Wasser (Verbrühungsgefahr). Von dieser muss bei der vorliegenden Evidenz-Lage explizit abgeraten werden.

31.5 Schlussfolgerung für die Praxis

Die Ärzte des Qualitätszirkels beschließen, das eigene Merkblatt mit folgender Zielsetzung zu ändern:

- ▲ Die Priorität der Kortison-Behandlung muss klar herausgestellt werden.
- ▲ Aufgrund des möglichen Nutzens wird die Exposition des Kindes zu kühler und

(subjektiv wohltuend empfundener) befeuchteter Luft angeraten.

- ▲ Aufgrund geänderter Nutzen-Risiko-Bewertung wird ausdrücklich von der Luftbefeuchtung durch erhitztes/kochendes Wasser abgeraten.

Die entsprechenden Passagen des alten und neuen Merkblattes sind in Tab. 31.2 zusammengestellt.

Literatur

- [1] Gesellschaft für Pädiatrische Pneumologie. Stenosierende Laryngotracheitis (Krupp) und Epiglottitis. <http://www.uni-duesseldorf.de/awmf/ll-na/026-006.htm> (01.12.06)
- [2] Scolnik D et al., Controlled delivery of high vs low humidity vs mist therapy for croup in emergency departments: a randomized controlled trial. JAMA. (2006), 295, 1274–1280
- [3] Neto GM et al., A randomized controlled trial of mist in the acute treatment of moderate croup. Acad Emerg Med. (2002), 9, 873–879
- [4] Bouchier D, Dawson KP, Fergusson DM, Humidification in viral croup: a controlled trial. Aust Paediatr J. (1984), 20, 289–291
- [5] Segal AO et al., Croup hospitalizations in Ontario: a 14-year time-series analysis. Pediatrics. (2005), 116, 51–55
- [6] Greally P et al., Children with croup presenting with scalds. BMJ. (1990), 301, 113

Tab 31.2: Vergleich der Therapie-Empfehlung für Eltern von Kindern mit Croup-Rezidiv

Text	Bei den ersten Krankheitszeichen versuchen Sie, Ihr Kind so gut wie möglich zu beruhigen.	
	alt Feuchten Sie die Atemluft gut an (feuchte Tücher auf die Heizkörper o.Ä.) oder setzen Sie das Kind ans offene Fenster (Kaltluft). Tritt keine Besserung ein, verabreichen Sie das verschriebene (Kortison-)Präparat.	neu Verabreichen Sie das verschriebene (Kortison-)Präparat. Feucht-kalte Luft (offenes Fenster) kann zusätzlich helfen. Verwenden Sie NIEMALS heißes/dampfendes Wasser in der Nähe des Kindes.
Grundlage	Lehrbuch AWMF-Leitlinien	Bewertung der neuen Literatur

32 Evidenzbasierte Medizin bei psychischen Erkrankungen

Michael M. Berner, Mathias Berger, Martin Härter

32.1 EbM in der Psychiatrie und Psychotherapie

Immer noch sind Psychiatrie, Psychosomatik und Psychotherapie Fachrichtungen, deren Therapieverfahren in der öffentlichen Meinung und in den Medien häufiger große Vorbehalte entgegengebracht werden. Der Bedarf an evidenzbasierten Vorgehensweisen in der Psychiatrie und Psychotherapie ist hoch: Zwei unabhängige Untersuchungen kamen 1995 und 1997 übereinstimmend zum Ergebnis, dass über ein Drittel aller Therapieentscheidungen in der Psychiatrie (noch) nicht durch überzeugende Evidenz aufgrund randomisierter kontrollierter Studien abgesichert ist [1; 2]. Psychiatrie und Psychotherapie richten sich seit vielen Jahren immer stärker an den Grundlagen der evidenzbasierten Medizin aus, wie z.B. auch neuere Lehrbücher oder der von der Fachgesellschaft beispielhaft beschrittene Weg der Leitlinienentwicklung zeigen [3; 4] (www.dgppn.de/leitlinien/leitlinien.html). Seit 2000 wurden zu den großen Indikationsbereichen des Fachgebietes (Schizophrenie, Affektive Störungen, Angst- und Ess-Störungen, Demenzen) entsprechende Leitlinien veröffentlicht, zuletzt die S3-Leitlinie zur Schizophrenie.

32.1.1 Depressive Erkrankungen

Psychische Erkrankungen, insbesondere depressive Störungen, die im Zentrum dieses Kapitels stehen, gehören zu den häufigsten Beratungsanlässen in der Patientenversor-

gung [5]. Die Erforschung der Diagnose- und Behandlungsmöglichkeiten hat in den letzten Jahren große Fortschritte gemacht. Hauptprobleme der Versorgung bei depressiven Erkrankungen betreffen vor allem das richtige und rechtzeitige Erkennen und die Ausrichtung der Behandlung an aktuellen Leitlinien. Darüber hinaus besteht im Gesundheitswesen Klärungsbedarf, wie eine abgestufte und vernetzte Versorgung zwischen haus- und fachärztlicher sowie psychotherapeutischer Behandlung organisiert sein sollte und was die Grundlage für die Indikationsstellung für ambulante und stationäre bzw. rehabilitative Behandlungsmaßnahmen darstellt. Auf Seite der Patienten bestehen nicht selten Vorbehalte gegenüber evidenzbasierten Therapieverfahren, wie Pharmako- oder Psychotherapie, die eine suffiziente Behandlung erschweren können [6–9].

Betrachtet man die Gesamtzahl aller depressiven Patienten, so zeigt sich, dass eine weit größere Anzahl in der primärärztlichen Versorgung behandelt wird als von psychiatrischen Fachärzten oder von ärztlichen bzw. psychologischen Psychotherapeuten. Dies kann bei schwierigeren Fragen bzgl. der Diagnosestellung (Art und Qualität der depressiven Störung, Komorbidität etc.) oder bei komplexen Therapieentscheidungen (z.B. Pharmako- vs. Psychotherapie vs. Kombinationsbehandlung) zu Problemen führen. Während Standard-Antidepressiva z.B. oft – den Ergebnissen kontrollierter Studien widersprechend – nur sehr niedrig dosiert werden, ist das am häufigsten verschriebene Präparat ein pflanzliches Produkt – das Hypericum-Extrakt (Johanniskraut). Ein klinisches

Beispiel soll anhand des von Sackett vorgeschlagenen fünfstufigen Handlungsalgorithmus verdeutlichen, wie praktisch Tätige mithilfe ausgewählter Methoden der EbM Therapieentscheidungen auf dem aktuellsten wissenschaftlichen Stand treffen können.

32.1.2 Leitlinien zu depressiven Erkrankungen

Die Spitzenverbände der Selbstverwaltungs-Körperschaften im Gesundheitswesen und das Ärztliche Zentrum für Qualität in der Medizin (ÄZQ) entwickelten 1996 ein Leitlinien-Qualitätsprogramm. Die Leitlinien-Clearingstelle des ÄZQ führte im Rahmen dieses Programms eine Überprüfung von Behandlungsleitlinien für zentrale Gebiete der medizinischen Versorgung durch [10] und veröffentlichte die Ergebnisse in sog. Leitlinien-Clearingberichten (www.leitlinien.de). Ein nationales Clearingprojekt zur Depression wurde 2001 vom ÄZQ initiiert. Hauptziel war die Identifikation und der Vergleich von evidenzbasierten, qualitativ hochwertigen deutschen und englischsprachigen Leitlinien zur Depression. Für den Zeitraum 1991–2001 wurden 21 Leitlinien (davon drei deutsche, s.u.) formal mit der Checkliste „Methodische Qualität von Leitlinien“ [11] bewertet. Die Inhalte der Leitlinien wurden bezüglich der Methodik der Leitlinienentwicklung, des Inhalts und Formats sowie der praktischen Anwendbarkeit beurteilt. Es zeigten sich deutliche Unterschiede in der Qualität bezüglich des Entwicklungsprozesses, der Verknüpfung der Empfehlungen mit der Evidenz und der Empfehlungen zur Implementierung. Der veröffentlichte Clearingbericht enthält relevante Eckpunkte und einen Gliederungsvorschlag für eine Nationale Leitlinie, um diese zu realisieren. In diesem Zusammenhang wurde auch ein Maßnahmenkatalog zur Weiterentwicklung und Implementierung formuliert [12].

Für Deutschland sind seit kurzem die überarbeitete Neuauflage der 1997 erstmals von der Arzneimittelkommission der deutschen Ärzteschaft für den hausärztlichen Bereich [13] erstellte Leitlinie und die bereits erwähnte, 2000 von der Deutschen Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde für den fachärztlichen Bereich [14] herausgegebene Leitlinie von Bedeutung.

Für die ambulante Versorgung, insbesondere die hausärztliche Versorgung, stehen seit 2003 die „Versorgungsleitlinien zur Diagnostik und Therapie depressiver Störungen in der hausärztlichen Versorgung“ zur Verfügung [15]. Diese Versorgungsleitlinien entstanden im Rahmen des Kompetenznetzes Depression, Suizidalität in Modellprojekten zum ambulanten und stationären Qualitätsmanagement. Obwohl in vielen Ländern zwischenzeitlich umfangreiche nationale Programme zur Entwicklung von Versorgungsleitlinien initiiert worden sind [16; 17], stößt die Implementierung von Leitlinien in der Praxis auf Schwierigkeiten. Unter anderem werden das Format und die Formulierung solcher Leitlinien kritisiert, außerdem wird gefordert, dass Versorgungsleitlinien enger am Praxisalltag zu orientieren sind und mit konkreten Anleitungen zur Umsetzung formuliert werden sollen [18; 19]. Eine Auswahl von internationalen, kontrollierten Studien im Bereich Depression und deren zentrale Ergebnisse hinsichtlich der Effekte einer Leitlinien-Implementierung sind in Tabelle 32.1 dargestellt.

In Deutschland wurden im Rahmen des vom BMBF geförderten Modellprojektes im Kompetenznetz Depression, Suizidalität in einer Interventions-Kontrollgruppen-Studie in drei Studienregionen (Südbaden, Aachen/Düren und München) die entwickelten Leitlinien für die ambulante Versorgung im Rahmen eines umfassenden Qualitätsmanagementsystems für depressive Erkrankungen, u.a. mit gezielten Fortbildungsmaßnahmen,

Tab. 32.1: Effekte einer Leitlinien-Implementierung auf die primärärztliche Versorgung depressiver Patienten (RCTs)

Quelle	Interventionen	Veränderungen der Prozess- und Ergebnisqualität
Worrall et al. (1999)/ Worrall et al. (2000) (Kanada)	Interventionsgruppe: 3-h-Training Inhalt: Erklärung der Leitlinien, Diskussion vorbereiteter Fälle und von Fällen der Ärzte; Zusammen- arbeit mit einem Psychiater Kontrollgruppe: nur Leitlinien vorgelegt	Prozess
		+ Vermehrtes Verschreiben von Antidepressiva, mehr kontinuierliche Behandlungen
		– weniger korrekte Diagnosen
		Ergebnis
		++ nach 6 und 18 Monaten stärkere syndromale Besserung der Depression
Thompson et al. (2000) (Großbritannien)	Seminare in Großgruppen (Leitlinienanwendung) Training in Kleingruppen (Videomaterial, Kleingruppendiskussionen, Rollenspiele) Ausbilder waren noch 9 Monate für Fragen verfügbar	Prozess
		+ Sensitivität bzgl. depressiver Symptome gebessert
		Ergebnis
		+ Nach 6 Wochen und nach 6 Monaten: Höherer Anteil an Patienten mit Verbesserungen der Depression und geringerer Anteil mit immer noch vorhandener Depression
		– nicht erkannte Fälle: nach 6 Wochen vermehrt immer noch mit Depression und nach 6 Monaten weniger verbessert
Baker et al. (2001) (Großbritannien)	Einsatz von Leitlinien Nach 6 Wochen: Identifikation von Hindernissen bei deren Umsetzung Auf Hindernisse zugeschnittene Fortbildung	Prozess
		++ Erfragen des Suizidrisikos verbessert
		+ Vermehrt Behandlung mit Antidepressiva oder Psychotherapie, angemessene Dosierung der Antidepressiva, Kontrolluntersuchungen, Behandlungsdauer verbessert
		– Suizidrisiko nach 3 Wochen weniger oft erneut erfragt
		Ergebnis
		++ Depressionswerte besser nach 16 Wochen

++ = signifikante Verbesserung

+ = nicht signifikante Verbesserung

– = nicht signifikante Verschlechterung

in haus- und nervenärztlichen Praxen implementiert und evaluiert [20]. Übersichtsarbeiten kommen zu dem Schluss, dass zu den bisher untersuchten Strategien zur Umsetzung von Trainingsmaßnahmen insgesamt in der Medizin [21] bzw. zur Verbesserung der Depressionsbehandlung [22] sowohl po-

sitive Studien als auch Studien ohne positive Effekte vorliegen. Komplexe Interventionsstrategien, die zudem Umsetzungsbarrieren (Behandler-, Patienten- und strukturelle Barrieren) berücksichtigen, sind am erfolgreichsten. Ziemlich wahrscheinlich kann eine erfolgreiche Umsetzung von Leitlinien

in die Praxis dann erreicht werden, wenn auch Strategien zur kontinuierlichen Patientenbegleitung („Case-Management“) und zur engen Vernetzung der verschiedenen Versorgungsebenen angewandt werden.

32.2 Anwendungsbeispiel

32.2.1 Problemdefinition

Eine Patientin ruft morgens die junge Assistenzärztin einer psychiatrischen Poliklinik an, die ihr am Vortag wegen einer Depression mittelgradiger Schwere nach – mit der Patientin gemeinsam vorgenommener – Abwägung der Vor- und Nachteile von Psycho- und Pharmakotherapie einen Serotonin-Wiederaufnahmehemmer verordnet hat. Sie ist verunsichert: Ihr Hausarzt habe ihr soeben empfohlen, anstelle des verschriebenen ein Johanniskraut-Präparat (*Hypericum*) einzunehmen. Diese Präparate hätten wesentlich weniger Nebenwirkungen und seien mindestens genauso wirksam. Die Assistenzärztin vereinbart einen erneuten Termin mit ihr für den Nachmittag. Ihr Oberarzt empfiehlt ihr, in den „entsprechenden Leitlinien“ nachzuschlagen. In der ihr zur Verfügung stehenden Online-Version der Leitlinie „Affektive Störungen“ der Deutschen

Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde sind jedoch Hypericum-Präparate überhaupt nicht erwähnt.

Die Ärztin möchte sich mit dieser Auskunft nicht zufriedengeben. Sie definiert ihr klinisches Problem in der Frage: „Sind Hypericum-Präparate in ihrer Wirkung bei mittelgradiger Depression den Standard-Antidepressiva vergleichbar?“ und begibt sich auf die Suche nach passender Literatur.

32.2.2 Literaturrecherche

Gute Lehrbücher bilden die Wissensbasis für die medizinische Tätigkeit in einem Fachbereich. Sie werden jedoch erfahrungsgemäß nur im Abstand von einigen Jahren überarbeitet. Bereits im Kapitel 20 dieses Lehrbuches wurde über verschiedene Informationsmedien und die Methodik der Literaturrecherche gesprochen. Tabelle 32.2 liefert eine Übersicht über die für das Fachgebiet der Psychiatrie und Psychotherapie relevanten Informationsmedien.

Die Möglichkeiten, schnell die für eine Fragestellung relevante Literatur zu finden, sind trotz deutlicher Verbesserungen in den letzten Jahren nach wie vor eingeschränkt.

Tab. 32.2: Für die Psychiatrie relevante Quellen der Literatursuche

Art der Quelle	Psychiatrierelevante Quellentitel
Primärliteraturdatenbank	MEDLINE, EMBASE PsychInfo (speziell psycholog. Literatur) Cochrane Database of Controlled Trials (auf CD-ROM)
Sekundärliteratur	Evidence-based Medicine, Journal Evidence-based Mental health, Journal Info Neurologie und Psychiatrie Best Evidence (CD-Rom)
Systematische Übersichtsarbeiten	z.B. aus Cochrane Database of Systematic Reviews (ca. 270 psychiatr. Titel)
Sonstige nationale Informationsquellen (ausgewählte Beispiele)	www.kompetenznetz-depression.de www.kompetenznetz-schizophrenie.de www.kompetenznetz-demenz.de www.leitlinien.de

Tab. 32.3: Psychiatrische Review-Gruppen der Cochrane Collaboration und bisherige Anzahl von erstellten systematischen Übersichtsarbeiten (Stand: August 2006)

Psychiatr. Review-Gruppe der Cochrane Collaboration	Editorische Basis	Anzahl Reviews
Depression, Anxiety and Neurosis	Auckland, New Zealand	41
Dementia and Cognitive Improvement	Oxford, GB	62
Drugs and Alcohol	Lazio, I	16
Developmental, Psychosocial and Learning Problems	Bristol, GB	17
Schizophrenia	Birmingham, GB	98

Einen wesentlichen Beitrag auf internationaler Ebene hierzu leistet die Cochrane Collaboration (siehe Kapitel 43). Sie hat sich zum Ziel gesetzt, systematische Übersichtsarbeiten herzustellen, zu verbreiten und aktuell zu halten. Auf psychiatrischem Fachgebiet existieren derzeit fünf Review-Gruppen, die in Tabelle 32.3 dargestellt sind. Sie stellen innerhalb ihres Themenbereiches nach strengen methodologischen Kriterien systematische Übersichtsarbeiten her und überarbeiten sie spätestens alle zwei Jahre wieder. Für die Psychiatrie gibt es hier zurzeit etwa 250 Übersichtsarbeiten. In der Cochrane Library wird auch eine Datenbank gepflegt, die diejenigen kontrollierten und randomisierten klinischen Studien listet, die bei der Erstellung der Übersichtsarbeiten identifiziert wurden (bspw. durch manuelle Suche) oder aus verschiedenen medizinischen Datenbanken stammen (z.B. MEDLINE, EMBASE). Das Schlagwort „**Psychiatry and Psychology**“ liefert hier Zugriff auf die Abstracts von etwa 8 000 kontrollierten Studien.

Tatsächlich kann eine Literaturrecherche im Einzelfall sehr zeitraubend sein. Im Idealfall ist die vorhandene Evidenz zur Diagnostik und Therapie einer Erkrankung in Form einer evidenzbasierten Leitlinie aufgearbeitet.

Die systematische Entwicklung von Leitlinien erfordert eine objektive und unbeeinflussbare Methodik, in der Ergebnisse der wissenschaftlichen Forschung (Prinzipien der Evidenzbasierten Medizin) sowie Expertenwissen (klinische Erfahrung) gleichermaßen berücksichtigt werden [23]. Idealerweise

enthalten sie Handlungsanweisungen zum konkreten Vorgehen bei diagnostischen und therapeutischen Maßnahmen. In den letzten Jahren sind weltweit eine große Zahl an Leitlinien und Behandlungsempfehlungen, u.a. auch für depressive Erkrankungen, entwickelt worden. Diese unterscheiden sich bzgl. ihrer Zielsetzung und ihres Umfangs, vor allem aber in ihrer methodischen Qualität [24; 25].

Im klinischen Beispiel fördert die Literaturrecherche in MEDLINE mit dem Suchparameter „**Hypericum**“ und „**review**“ eine aktuelle systematische Übersichtsarbeit mit einer Auswertung von insgesamt 39 randomisierten kontrollierten Studien zutage [26].

Die aufgefundene Übersichtsarbeit beinhaltet eine Meta-Analyse, eine statistische Zusammenfassung der Ergebnisse unterschiedlicher Studien. Auch Meta-Analysen sind nicht vor falschen Schlussfolgerungen gefeit und müssen kritisch betrachtet werden. Verzerrungen können zum Beispiel dadurch entstehen, dass nicht alle vorhandenen relevanten Studien in die Meta-Analyse einbezogen wurden.

Hätte die Ärztin beispielsweise den Versuch unternommen, die in der Übersichtsarbeit dargestellten Arbeiten direkt aufzufinden, hätte sie in der Datenbank MEDLINE nur insgesamt 18 der 39 in der Übersichtsarbeit gepoolten Studien aufgefunden (Retrieval-Bias).

Die weiteren verwendeten Datenbanken standen – aus Kostengründen (EMBASE) oder weil sie nur für einen sehr eingeschränkten

Benutzerkreis gedacht sind (PHYTODOC) – nicht zur Verfügung. Diese Schwierigkeiten gelten insbesondere natürlich für Studien, die – oftmals wegen negativer Ergebnisse – niemals publiziert wurden (Publication-Bias) [27; 28]. Ebenso lässt sich als weitere systematische Abweichung der „English language-Bias“ demonstrieren: Hätte man sich in der Übersichtsarbeit, wie oft üblich, rein auf englischsprachige Publikationen beschränkt, wäre keine einzige der vor 1999 publizierten Studien aufgefunden worden.

32.2.3 Kritische Bewertung der aufgefundenen Evidenz

Der zentrale Schritt der Anwendung der Evidence-based Medicine ist die kritische Beurteilung der aufgefundenen Evidenz hinsichtlich Validität und Relevanz.

Die systematische Übersichtsarbeit unseres Beispiels – es handelt sich dabei um eine für die Cochrane Library vorgesehene Übersichtsarbeit – gibt einen Überblick über die vorliegenden randomisierten Studien, enthält einen ausführlichen methodischen Abschnitt, kommentiert und bewertet die einzelnen Studien nach klaren Kriterien. Sie wird somit von der Ärztin als valide bewertet. Die unterschiedlichen diagnostischen und psychometrischen Messmethoden in den einzelnen Studien werden von den Autoren detailliert kommentiert.

Die Arbeit kommt zu dem Schluss, dass in den Studien, die sich auf Patienten mit Major Depression beschränken, ein geringer Vorteil von *Hypericum* gegenüber Placebo besteht. Die insgesamt 14 Studien mit Vergleichen von Standard-Antidepressiva mit insgesamt 2238 Patienten weisen auf eine ähnliche Effektivität hin. Die Suche und Literaturbeschaffung dauerte insgesamt 40 Minuten. Kurze Zeit darauf gibt der Oberarzt seiner Assistentin eine aktuelle Praxisleitlinie für depressive Störungen [13], die die Ärztin sich

auch über das Internet hätte beschaffen können (<http://www.akdae.de/35/index.html>). In ihr gibt es einen ausführlichen Abschnitt über Phytotherapeutika mit Aufarbeitung der wissenschaftlichen Evidenz. Der Abschnitt enthält die auf der Basis systematischer Übersichtsarbeiten ausgesprochene Empfehlung, dass ein Therapieversuch mit Johanniskrautpräparaten bei „Patienten mit eher leichter bis mittelgradiger Depression“ in Erwägung gezogen werden könnte (angezeigt durch das Symbol „↑↑“).

32.2.4 Integration der aufgefundenen Evidenz in die klinische Arbeit

An dieser Stelle kommt die klinische Erfahrung zum Tragen: Treffen die jeweiligen Studienergebnisse auf den Patienten zu, der zur Fragestellung Anlass gegeben hat? Ist der Patient so verschieden, dass er nicht in die Studien aufgenommen worden wäre? Weiterhin ist es aber auf der subjektiven Ebene sehr wichtig, dass Arzt und Patient im Sinne der partizipativen Entscheidungsfindung („Shared decision making“) im Gespräch bleiben und sich über ihre Wertvorstellungen und Ziele der Behandlung klar werden.

Am Nachmittag schildert die Ärztin ihrer Patientin kurz die Inhalte der Übersichtsarbeit und erläutert die Gründe, warum sie bei ihrer Therapieempfehlung bleibt. Sie integriert so die Ergebnisse aktueller Forschung in ihr klinisches Handeln. Sie weist ausdrücklich auf die Inhalte der Leitlinie hin und beteiligt die Patientin an der Entscheidungsfindung. Nach einer erneuten Aufklärung über die möglichen unerwünschten Effekte des Serotonin-Wiederaufnahmehemmers ist die Patientin beruhigt und erklärt, sie werde das verordnete Medikament einnehmen.

32.2.5 Evaluation der ärztlichen Leistung

Am darauf folgenden Tag könnte nun die Assistenzärztin im Sinne einer Evaluation der eigenen Leistung ihrem Oberarzt die Ergebnisse der Recherche präsentieren. Wenn dieser sich dem Konzept einer Evidenzbasierten Medizin gegenüber interessiert und abgeschlossen zeigt, könnte er die Ärztin um eine Zusammenstellung im Rahmen eines 5-Minuten-Vortrages für die tägliche Dienstbesprechung bitten und könnte die Ärztin, die sich in ihrer Verschreibung an einer aktuellen Leitlinie orientiert hat, mit einer weiteren Recherche betrauen, nämlich der Frage, inwieweit sich eine Verbesserung der täglichen Arbeit durch leitlinienorientiertes Vorgehen erzielen lässt.

32.3 Schlussfolgerungen

Evidenzbasierte Medizin ist eine Methode, die den Austausch von Informationen im therapeutischen Team fördern kann. Durch die immanente kontinuierliche Evaluation der eigenen Leistung und der Effektivität von angewandten Maßnahmen kann sie in einer Zeit der knapper werdenden finanziellen Ressourcen ein entscheidendes Instrument der Qualitätssicherung darstellen. Sie verbessert den Informationsstand des klinisch Tätigen durch kontinuierliche, selbstgesteuerte Weiterbildung und liefert methodische Beurteilungskriterien zur Einschätzung neuer Behandlungsmethoden und Ergebnisqualitäten.

Zunächst ist eine Realisierung für den klinisch Tätigen nicht ohne zeitlichen Mehraufwand realisierbar, doch ist dieser Aufwand für Recherchen angesichts der zunehmenden Verfügbarkeit elektronischer Datenbanken am Arbeitsplatz und der häufiger werdenden Möglichkeit der Volltextsuche in diesen Datenbanken begrenzt. Außerdem wiederholen sich Problemstellungen, sodass

die einmal recherchierte und aktualisierte Problemstellung beim nächsten Mal bereits unmittelbar verfügbar ist. Zuletzt wird es von einem Psychiater als Spezialisten seiner Fachrichtung erwartet, einen Überblick über den „State of the Art“ seines Faches, also die neuesten Forschungsergebnisse, zu haben.

Die Entwicklung und Verwendung von Leitlinien kann hier eine Hilfestellung sein. Leitlinien sind jedoch nur dann sinnvoll, wenn sie einer wissenschaftlichen Überprüfung unterzogen werden. Die Entwicklung eines Implementierungs- und Evaluationskonzeptes ist daher eine *conditio sine qua non* für eine qualitativ hochwertige, evidenzbasierte Leitlinie. Aktuelle, evidenzbasierte Leitlinien stellen die derzeit beste Möglichkeit für Leistungserbringer und -empfänger dar, um nachvollziehbare und überprüfbare Maßnahmen einzuleiten. Letztendlich müssen aber im konkreten Einzelfall im Sinne einer partizipativen Entscheidungsfindung sowohl ärztliche als auch patientenrelevante Faktoren wie Wissen, Erfahrungen, Präferenzen sowie soziokulturelle Aspekte berücksichtigt werden.

Einer weiteren Schwierigkeit müssen sich diese Leitlinienformulierungen ebenfalls stellen: Sehr wahrscheinlich besteht ein Ungleichgewicht hinsichtlich der Verfügbarkeit von guten Studien in unterschiedlichen Bereichen der Behandlung psychischer Erkrankungen. Auf diese Weise fließt das vorherige wissenschaftliche Investment indirekt in die Empfehlungsstärke mit ein. Ein großes Problem in der Behandlung psychischer Erkrankungen stellt beispielsweise das relative Fehlen von Langzeitstudien dar, obwohl die meisten Behandlungen über Monate und Jahre dauern. Eine kürzlich mitgeteilte Auswertung der Cochrane-„Depression Anxiety and Neurosis Group“ ergab, dass lediglich 8% aller relevanten Studien im Bereich Pharmakotherapie der Depression eine Dauer von mehr als 24 Wochen aufwiesen. Ebenso stellt die Repräsentation der Primärpraxis

mit nur 25% aller Studien (dagegen 75% im stationären Rahmen) eine Umkehrung gegenüber den klinischen Verhältnissen dar. Ein anderer Punkt, der in letzter Zeit in der Psychiatrie verstärkt Beachtung fand, war die Publikationspolitik pharmazeutischer Firmen. Im Jahr 2004 erregte eine Meta-Analyse [29] großes Aufsehen, die in einer Auswertung von 11 Studien zeigte, dass publizierte – und somit deutlich besser verfügbare – Daten zum Gebrauch neuerer Antidepressiva bei Kindern und Jugendlichen generell positiver für das jeweilige Medikament ausfielen als die unpublizierten. In der Meta-Analyse blieb lediglich noch ein positiver Effekt für Fluoxetin übrig. Des Weiteren fanden sich in dieser Analyse Hinweise für ein erhöhtes Suizidrisiko in dieser Altersgruppe durch selektive Serotonin-Wiederaufnahmehemmer (SSRI) und atypische Antidepressiva.

Weiter ist es zu betrachten, dass es medizinische Maßnahmen gibt, die in der Praxis ihre Wirksamkeit gezeigt haben (z.B. die Elektrokrampftherapie – EKT), ohne durch klinische Studien hinreichend überprüft worden zu sein. Außerdem gibt es Fachbereiche (zu denen auch in besonderem Maße die Psychiatrie und Psychotherapie gehört), in denen es sich häufig aus ethischen Gründen verbietet, randomisierte Studien durchzuführen. So wird es weiterhin viele medizinische Entscheidungen geben, die auf Erfahrung, Analogieschlüssen und Extrapolation beruhen müssen. Diese Vorgehensweisen schließt Evidenzbasierte Medizin nicht aus. Sie verlangt nur die Kenntnis der zurzeit besten verfügbaren Evidenz als Hilfsmittel zur individuellen, am Patienten orientierten Entscheidungsfindung.

Literatur

- [1] Geddes J et al., Inpatient psychiatric treatment is evidence-based. *Qual Health Care.* (1996), 5, 215–217
- [2] Summers A, Kehoe RF, Is psychiatric treatment evidence-based? *Lancet.* (1996), 347, 409–410
- [3] Berger M (Ed.) (2004) *Psychische Erkrankungen – Klinik und Praxis.* Urban und Fischer, München
- [4] Berner MM, Ruther A, Stieglitz RD, Berger M, Evidenzbasierte Psychiatrie – ein Weg zu einer rationaleren Psychiatrie? *Nervenarzt.* (2000), 71, 173–180
- [5] Murray C, Lopez AD (1997) Global and regional descriptive epidemiology of disability: incidence, prevalence, health expectancies and years lived with disability. In: Murray C, Lopez AD, *The global burden of disease.*, 201–246. Harvard Univ. Pr., Harvard
- [6] Bermejo I, Lohmann A, Berger M, Härter M, Barrieren und Unterstützungsbedarf in der hausärztlichen Versorgung depressiver Patienten. *Z Arztl Fortbild Qualitätssich.* (2002), 96, 605–613
- [7] Donoghue JM, Tylee A, The treatment of depression: prescribing patterns of antidepressants in primary care in the UK. *Br J Psychiatry.* (1996), 168, 164–168
- [8] Paykel ES et al., The Defeat Depression Campaign: psychiatry in the public arena. *Am J Psychiatry.* (1997), 154, 59–65
- [9] Schneider F et al., Insufficient depression treatment in outpatient settings. *Ger Med Sci.* (2004), 2, DOC01 (20040226)
- [10] Ollenschläger G et al., The German Guidelines Clearing House (GGC) – Rationale, Aims and Results. *Proc R Coll Physicians Edinb.* (2001), 31 (Suppl. 9), S59–S64
- [11] Ärztliche Zentralstelle Qualitätssicherung (ÄZQ), Checkliste „Methodische Qualität von Leitlinien“. *Dt Arztebl.* (2000), 97, A–1170–2
- [12] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ). Leitlinien-Clearingbericht „Depression“. <http://www.leitlinien.de/clearingverfahren/clearingberichte/depression/00depression/view> (13.09.05)
- [13] Arzneimittelkommission der deutschen Ärzteschaft (AkdÄ). Depression. 2. Auflage. <http://www.akdae.de/35/67-Depression-2006-2Auflage.pdf> (14.11.06)

- [14] Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde (DGPPN) (2000) Behandlungsleitlinie Affektive Erkrankungen. Steinkopff, Darmstadt
- [15] Härter M et al., Versorgungsleitlinien zur Diagnostik und Therapie depressiver Störungen in der hausärztlichen Praxis. *Z Arztl Fortbild Qualitätssich.* (2003), 97 Suppl 4, 16–35
- [16] Burgers JS et al., Towards evidence-based clinical practice: an international survey of 18 clinical guideline programs. *Int J Qual Health Care.* (2003), 15, 31–45
- [17] Härter M et al., Vernetzte Versorgung depressiver Patienten – Entwicklung eines Behandlungskorridors und Implementierung evidenzbasierter Behandlungsempfehlungen. *Z Allg Med.* (2003), 78, 231–238
- [18] Davis D et al., Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA.* (1999), 282, 867–874
- [19] Michie S, Johnston M, Changing clinical behaviour by making guidelines specific. *BMJ.* (2004), 328, 343–345
- [20] Härter M et al., Umfassendes ambulantes Qualitätsmanagement in der Versorgung depressiver Patienten. *Z Arztl Fortbild Qualitätssich.* (2003), 97 Suppl 4, 9–15
- [21] Grimshaw JM et al., Changing provider behavior: an overview of systematic reviews of interventions. *Med Care.* (2001), 39, II2–45
- [22] von Korff M, Goldberg D, Improving outcomes in depression. *BMJ.* (2001), 323, 948–949
- [23] Kopp I, Encke A, Lorenz W, Leitlinien als Instrument der Qualitätssicherung in der Medizin. Das Leitlinienprogramm der Arbeitsgemeinschaft Wissenschaftlicher Medizinischer Fachgesellschaften (AWMF). Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz. (2002), 45, 223–233
- [24] Helou A, Perleth M, Schwartz FW, Prioritätensetzung bei der Entwicklung medizinischer Leitlinien. Teil 1: Kriterien, Verfahren und Akteure: eine methodische Bestandsaufnahme internationaler Erfahrungen. *Z Arztl Fortbild Qualitätssich.* (2000), 94, 53–60
- [25] Helou A et al., Methodische Standards der Entwicklung evidenz-basierter Leitlinien in Deutschland. Konsens zwischen Wissenschaft, Selbstverwaltung und Praxis. *Z Arztl Fortbild Qualitätssich.* (2000), 94, 330–339
- [26] Linde K et al., St John's wort for depression: meta-analysis of randomised controlled trials. *Br J Psychiatry.* (2005), 186, 99–107
- [27] Dickersin K, The existence of publication bias and risk factors for its occurrence. *JAMA.* (1990), 263, 1385–1389
- [28] Easterbrook PJ et al., Publication bias in clinical research. *Lancet.* (1991), 337, 867–872
- [29] Whittington CJ et al., Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet.* (2004), 363, 1341–1345

33 Evidenzbasierte Medizin in der Sozialmedizin: Beispiel eines Grundsatzgutachtens zur Vertebroplastie und Kyphoplastie

Johannes Giehl, Christoph Kreck

33.1 Einführung

Evidenzbasierte Medizin im ärztlichen Alltag bedeutet für Sozialmediziner in den Medizinischen Diensten der Krankenversicherung (MDK), dass bei Beurteilungen im Einzelfall wie zu allgemeinen Fragen der Versorgung Methoden der EbM mit spezifischen Kenntnissen der Sozialrechtsprechung, der Versorgungsstrukturen und den Anforderungen der Krankenkassen und ihrer Verbände verknüpft werden. Das Spannungsfeld zwischen fachlich-medizinischer Unabhängigkeit und Einbindung als Berater in das GKV-System wird dabei deutlich [1]. Die Medizinischen Dienste haben sozialmedizinische Expertengruppen gebildet, die u. a. die Einheitlichkeit der Begutachtung, die Bearbeitung von speziellen Fragestellungen und auch die vorausschauende Marktbeobachtung zum Ziel haben. Bei der Beratung durch die Medizinischen Dienste muss die Qualität dem Stand der wissenschaftlichen Methodik entsprechen. Dies ist für die Beurteilung von Methoden im Einzelfall in der „Begutachtungsanleitung NUB 2002“ [2] operationalisiert. Für die Grundsatzberatung bestehen in einer Projektgruppe erarbeitete methodische Empfehlungen [3]. Dies ist die Grundlage einheitlicher Qualität der MDK-Produkte über alle Versorgungssektoren und Methoden hinweg; die Resultate sind im Falle von Meinungsverschiedenheiten zwischen Leistungsträgern, Leistungserbringern und Öffentlichkeit belastbar. Als Beispiel für eine Beratung zu innovativen Methoden stellen wir die Grundsatzbegutachtung zu Vertebroplastie (Vp) und Kyphoplastie (Kp) [4] vor.

Hier erfolgte die Beratung nicht in der Ideal-situation im Vorfeld, sondern nach einer teilweisen Etablierung im Gesundheitsmarkt.

33.2 Sozialmedizinischer Hintergrund und Versorgungsproblem

Osteoporose ist eine Erkrankung des höheren und sehr hohen Alters. Die Prävalenz osteoporotischer Wirbeldeformitäten wird aufgrund einer Querschnittsuntersuchung mit 10,2–18,7% in Westeuropa angenommen und in Deutschland werden jährlich mehr als 200 000 neue Frakturen geschätzt [5]. Sozialmedizinisch bedeutsam ist außerdem, dass etwa ein Siebtel der Patienten chronische Schmerzen beklagt und in seiner Mobilität zunehmend beeinträchtigt ist. Zeigt die meist erfolgreiche konservative Frakturbehandlung (Schmerzmedikation, Physiotherapie, Orthesen) kein zufriedenstellendes Ergebnis, erfolgt konventionell eine operative Wirbelfusion mit häufigen Komplikationen. Seit einigen Jahren werden stattdessen auch Vp und Kp bei schmerzhaften osteoporotischen Wirbeldeformitäten, ob akut oder chronisch eingetreten, durchgeführt. Beides sind minimal-invasive Eingriffe, bei denen eine Zementaussteifung der frakturierten Wirbelkörper erfolgt. Diese plausibel erscheinende Maßnahme (Behandlung am pathologischen Wirbelkörper selbst statt Versteifung mehrerer Wirbel) wird kontrovers diskutiert, weil die Wirksamkeit der Verfahren bisher nicht belegt ist und die Indikationsstellung und die Bedeutung schwe-

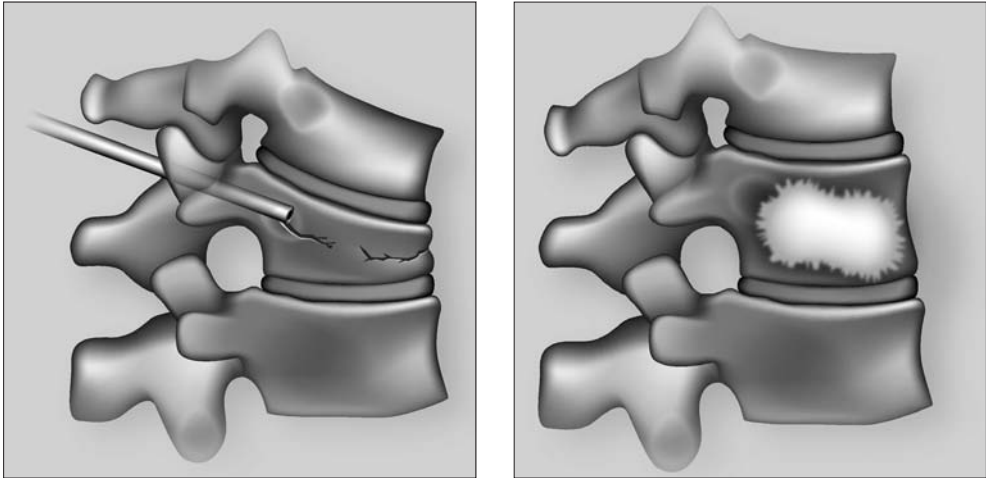


Abb. 33.1: Schematische Darstellung einer osteoporotischen Wirbelkörperfraktur mit transpedikulär platzierter Kanüle (links), Wirbelkörperaugmentation mittels Knochenzement (rechts)

rer Komplikationen noch nicht geklärt sind (s. Abb. 33.1).

Die **Operationstechnik** der Verfahren ist in zahlreichen Übersichtsartikeln beschrieben. Bei der Vp wird – meist in Lokalanästhesie – eine Kanüle unter Bildwandlerkontrolle in den Wirbelkörper vorgetrieben und dünnflüssiger Knochenzement unter hohem Druck injiziert, und zwar solange, bis die Wirbelfüllung nicht mehr zu erhöhen ist oder Zement die Wirbelumgebung austritt. Bei der Kp erfolgt – meist in Vollnarkose – nach der Platzierung der Kanüle zunächst die Einführung eines Ballons. Die Aufblähung des Ballons schafft ein Kavum (und eventuell eine Wirbelkörperaufrichtung), in das Knochenzement mit niedrigem Druck instilliert wird (Wirbelkörperaussteifung). Menge und Art des Knochenzementes sind bei beiden Verfahren nicht standardisiert; mehrere Ze-

mente sind für die Wirbelaugmentation CE-zertifiziert.

33.3 Auftrag und Arbeitsgruppe zur Gutachtenerstellung

Ein Krankenkassen-Bundesverband erteilte den Auftrag, ein Grundsatzgutachten zum Nutzen, zur Notwendigkeit und zur Wirtschaftlichkeit der beiden Verfahren zu erstellen. Zusätzlich sollte geklärt werden, ob diese Verfahren ambulant durchgeführt werden können. Die Fragen wurden spezifiziert (vgl. Tab. 33.1).

Wegen der komplexen Thematik der Wirbelbrüche bei Osteoporose wurde eine interdisziplinäre MDK-Arbeitsgruppe aus Orthopäden, spezialisierten Wirbelsäulenchirurgen, Neurochirurgen und Experten für

Tab. 33.1: Beispiele für spezifizierte Fragen

Besteht Evidenz der Wirksamkeit der Vp und der Kp?
Inwieweit sind sie kurz- und langfristig sicher?
Sind Vp und Kp mit zureichender Qualität ambulant erbringbar?
Sollen Vp und Kp Bestandteil der Versorgung in der gesetzlichen Krankenversicherung sein, und gegebenenfalls wie?

EbM-Methodik und für Medizinprodukte gebildet. Die Systemkompetenz der Gruppe wurde durch Erfahrungen aus verschiedenen Aufgabenfeldern der Medizinischen Dienste (ambulant – stationär, Einzelfall- bzw. Grundsatzberatung) sichergestellt.

Zunächst wurde der aktuelle Kontext zusammengetragen, u.a.:

- ▲ Häufige Kassenanfragen zur evtl. Indikation und zur ambulanten Abrechnung (derzeit keine EbM-Leistungen),
- ▲ Inhomogene Anwendungsfrequenz der Verfahren in orthopädischen und neurochirurgischen Kliniken, offensichtlich bedingt durch unterschiedliche Präferenzen,
- ▲ Zunehmende Verbreitung, auch in nicht spezialisierten Fachkliniken.

33.4 Arbeitsplanung und Literaturrecherche

Das Gutachten wurde nach den MDK-internen Empfehlungen geplant. Die Literaturrecherche wurde auf Ergebnisstudien, Komplikationen und technische Mitteilungen abgestellt. Um eine umfassende Aggregation des verfügbaren Wissens zu erhalten, wurde eine systematische Literaturrecherche in 28 Datenbanken vorgenommen. Es wurden Einschränkungen bezüglich des Publikationsjahres (ab 1987) und des Dokumententyps (RCTs, Controlled Clinical Trials, Clinical Trial, Meta-Analysis, Multicenter Study, Validation Studies) vorgenommen. Die wesentlichen Arbeiten wurden in MEDLINE, EMBASE, den HTA-Datenbanken und in der Cochrane Library gefunden. Ergänzend wurden Materialien von Herstellern sowie Informationen, die bei Kongressen erhältlich waren, hinzugezogen und die Literaturverzeichnisse der eingeschlossenen und relevanten Publikationen geprüft. Darüber hinaus erfolgte eine Umfeldrecherche in Google und auf den Internetseiten bekannter Hersteller. Auch

Auskünfte klinischer Wirbelsäulenchirurgen wurden herangezogen.

Es waren insgesamt 235 Treffer zu verzeichnen, wobei 8 HTAs zur Vp oder Kp aus den Jahren 2000 bis 2003 enthalten waren. Nach Durchsicht aller Abstracts wurde bei 108 Volltexten geprüft, ob es sich um Ergebnisstudien oder um Artikel zur Methodentechnik einschließlich Komplikationen handelte. Schließlich standen 12 Fallserien zur Vp und 5 zur Kp sowie je eine Vergleichsstudie zu Vp bzw. Kp zur Analyse zur Verfügung; die Ausschlussgründe für die übrigen Publikationen (Berücksichtigung bereits in den HTAs, Übersichtsartikel, Kommentare etc.) wurden einzeln dokumentiert.

33.5 Darstellung der Evidenz

8 HTAs/Reviews gründeten sich auf 64 zum Teil identische Studien (Fallserien). 6 dieser Berichte befassten sich mit der Vp; durchgehend wurden Erfolge bei Schmerzen und Mobilität in 84–100% berichtet. 2 Arbeiten untersuchten die Kp; die Erfolgsraten waren gleichermaßen hoch, eine Rekonfiguration der gebrochenen Wirbel war jedoch nicht augenfällig. Wegen der kleinen Fallzahlen der durchweg unkontrollierten Studien und der kurzen Beobachtungszeiten bewerteten 7 der 8 Berichte Wirksamkeit und Sicherheit der beiden Methoden als nicht belegt, sie empfahlen die Anwendung nur in kontrollierten klinischen Studien. Das HTA des britischen National Institute for Health and Clinical Excellence (NICE) von 2003 zur Vp [6] formulierte eine positive Empfehlung bei therapieresistenten Schmerzen, wobei allerdings „kompetente“ Ärzte und eine qualifizierte Klinikinfrastruktur vorliegen sollten.

Bei den insgesamt 13 Studien zur Vp handelte es sich mit einer Ausnahme um Fallserien. In Tabelle 33.2 sollen beispielhaft wichtige Daten aus der einzigen Vergleichsstudie [8] dargestellt werden. Die Autoren be-

Tab. 33.2: Vertebroplastie: Daten aus [7]

Einschluss: 79 Pat. mit osteoporotischen Wirbelfrakturen, retrospektiv
Intervention: Vertebroplastie früh nach der Fraktur (1–6 Wochen Schmerzdauer)
Kontrolle: 24 die Vp ablehnende Pat., Verum: 55 Pat./71 Wirbel
Stationäre Verweildauer bei OP im Mittel 9 Tage, bei konservativer Therapie 15 Tage
Beobachtungszeit 215 (57–399) Tage
Nach 24 Std. Schmerzlinderung bei 53% der Pat. mit OP, bei 29% Funktionsverbesserung
Nach 6 Wochen, 6 und 12 Monaten keine Unterschiede der Verum- und Kontrollgruppe
Komplikationen: 1 Querfortsatzfraktur und 2 Psoashämatome

Tab. 33.3: Kyphoplastie: Daten aus [8]

Einschluss: 60 Pat. mit osteoporotischen Wirbelfrakturen, prospektiv
Intervention: Kyphoplastie mindestens ein Jahr nach Fraktur/anhaltende Schmerzen
Kontrolle: 20 die Kp ablehnende Pat., Verum: 40 Pat./72 Wirbel
Beobachtungszeit: mind. 6 Monate
Schmerzbesserung: 26 auf 44 Punkte VAS (Verum), 33 auf 35 Punkte VAS (Kontrollen)
Kyphosewinkel: 8,7° auf 8,3° (Verum), 7,1° auf 8,3° (Kontrollen)
Wirbelkörperhöhe: 59 mm auf 65 mm (Verum), 63 mm auf 58 mm (Kontrollen)
Aktivitätsscore: 43,8 auf 54,4 (Verum), 39,8 auf 43,8 (Kontrollen)
Zementleckagen: 9,7% der operierten Wirbel

tonen die Wirksamkeit der Vp hinsichtlich Schmerz- und Funktionsverbesserung bei frischen Frakturen. Da die weitaus meisten osteoporotischen Frakturen nach wenigen Wochen konservativer Therapie problemlos ausheilen, bleibt offen, ob nicht zahlreiche Patienten unnötig operiert wurden. Die Wirksamkeit der Vp bei konservativ-therapieresistenten Patienten und die adäquate Operationsindikation bleiben ungeklärt, obgleich diese Studie (Evidenzklasse III der Agency for Healthcare Research and Quality/AHRQ) methodisch die höchstrangige war.

Bei den **Studien zur Kp** handelte es sich um 5 Fallserien und eine Vergleichsstudie [8], deren Ergebnisse beispielhaft dargestellt werden sollen (s. Tab. 33.3). Die Autoren resümieren eine positive Schmerz- und Aktivitätsbeeinflussung sowie eine Rekonfiguration der frakturierten Wirbel und eine protektive Wirkung gegen weitere Wirbelbrüche. In

dieser Studie (Evidenzklasse III nach AHRQ), die methodisch die höchstrangige war, ist der Selektions-Bias nicht ausgeräumt, die statistische Zuverlässigkeit der Ergebnisunterschiede und die langfristige Sicherheit der Kp sind nicht belegt.

Bezüglich **Komplikationen** wurden alle Ergebnisstudien und die Übersichten ausgewertet. Bei der Vp fanden sich perivertebraler Zementaustritt in 20–72% der behandelten Wirbel, Myelon- und Ösophagussschäden sowie Embolien bei venöser Zementverschleppung in Einzelfällen, Fettembolien (selten), Querfortsatz- und Rippenfrakturen (selten) und direkte Myelonverletzungen (Einzelfälle); ein häufigeres Auftreten von Folgefrakturen ist nicht belegt. Bei der Kp bestehen die gleichen Komplikationen, extravertebrale Zementaustritte sind jedoch seltener.

Alle Publikationen empfehlen daher eine kontrollierbare Zementinjektion unter Bild-

wandlerkontrolle, unverzügliche wirbelsäulen-chirurgische Interventionsmöglichkeit und die Verfügbarkeit einer Intensivstation. Mindestmengen für Operateure wurden nicht thematisiert; eine Firma bietet Kurse zur technischen Durchführung an. Mehrheitlich wurden die Studienpopulationen stationär behandelt.

33.6 Synthese des Wissens

Vp und Kp sind im Vergleich zu Wirbelfusionen weniger invasive Eingriffe, die aber technisch diffizil sind und schwerwiegende Komplikationen nach sich ziehen können.

Wirksamkeit: Schmerzlinderung nach Vp wird in 70,9–90,0% berichtet, nach Kp in 87,0–100,0%. Verbesserungen in den Alltagsaktivitäten werden nach Vp in 52,5–78,6% angegeben, nach Kp ist eine Verbesserung des Oswestry Disability Index (ODI) in einer Studie von 43,8 auf 54,4% angegeben. Der Zusatznutzen der Kp (Wirbelrekonfiguration) ist nicht belegt.

Sicherheit: Die kumulierten Komplikationen ergeben (exklusive Leckagen) für die Vp 5,9%, für die Kp 2,3%. Zementleckagen ereignen sich in 56,4% bzw. in 7,3% der Fälle. Schwerwiegende neurologische Komplikationen sind sehr selten. Nachteilige Effekte der Zementimplantation sind nach bis zu 5-jähriger Beobachtungszeit bei der Vp nicht festgestellt; bei der Kp ist das Follow-up noch zu kurz.

Langzeiteffekte: Nach Vp wurde eine Nachhaltigkeit positiver Wirkungen über 5 Jahre berichtet, für die Kp ist die Beobachtungszeit weniger als ein Jahr. Nach Vp ist eine erhöhte Frakturrate benachbarter Wirbel nicht nachzuweisen, hinsichtlich der Kp sind die Beobachtungszeiten zu kurz.

Fazit: Die Evidenz für die Wirksamkeit, die Sicherheit und für eine Bewertung des langfristigen Nutzens beider Verfahren ist insgesamt schwach. Dies entspricht der Situation bei vielen Interventionen unter Einsatz von Medizinprodukten, obwohl hierfür Studien höheren Evidenzniveaus realisierbar wären.

33.7 Sozialmedizinische Empfehlung

Vp und Kp scheinen ein Wirksamkeitspotenzial zur Schmerzverringerung und Mobilitätsverbesserung bei anhaltend schmerzhaften osteoporotischen Wirbelkörperfrakturen zu haben. Die Vp erscheint sicher, bei der Kp fehlen hinreichend langfristige Beobachtungen. Die vorliegenden Studien niedriger Evidenzklassen lassen eine eindeutigere Aussage nicht zu. Der optimale Zeitpunkt der Indikationsstellung ist wohl 4–6 Wochen nach Fraktureintritt bei unzureichend gebesserten Schmerzen trotz konservativer Therapie.

Es sollte geprüft werden, ob Anforderungen an die fachliche Qualifikation der Behandler und die Strukturqualität der Einrichtungen festgelegt werden können, da die Vp personelle und institutionelle Voraussetzungen erfordert, die erst eine hohe Ergebnisqualität erwarten lassen. Der zusätzliche Nutzen der Kp wäre noch zu belegen.

Die Anwendung der EbM-Methoden führt zu einer begründeten und transparenten Beurteilung des Stands der Erkenntnisse zur Wirksamkeit und den Risiken beider Verfahren. Es werden jedoch auch Limitierungen deutlich. Anders als bei Arzneimitteln besteht bei Methoden – besonders im stationären Bereich – keine klare Differenzierung zwischen klinischer Erprobung, Einführung in die GKV-Versorgung und Weiterentwicklung. So fehlen bei der Einführung von Methoden in die Versorgung häufig aussagekräftige Studien zu Wirksamkeit und Sicherheit. Nachdem die Methode erfolgreich

etabliert wurde, ist das Interesse der Anwender gering. Studien zur Beantwortung grundsätzlicher Fragen aufzulegen. Dies führt zu teilweise erheblicher Diskrepanz zwischen der Verbreitung in der Praxis und bestehender Evidenz. Eine evidenzbasierte Bewertung bereits etablierter Methoden kann mit der Dynamik der Verbreitung im Konflikt stehen. Die Bewertung kann aber ggf. dazu beitragen, dass Methoden zukünftig auf Evidenzbasis eingeführt oder weiterentwickelt werden oder dass auch Fehlentwicklungen revidiert werden.

Bei der Formulierung konkreter realisierbarer Empfehlungen bei der trotz unzureichender Evaluation schon fortgeschrittenen Etablierung der Verfahren ist zu bedenken, dass z.B. ein Ausschluss (sektoral oder für alle Sektoren) durch den G-BA kaum zu erwarten ist. Ein Ausschluss ist in Anbetracht der internationalen Entwicklung (z.B. [9]) und aus sozialmedizinischer Sicht auch nicht anzustreben. Die Methode sollte mit festgelegten Qualitätsstandards und bei weiterer Evaluation für therapierefraktäre Patienten zur Verfügung stehen. Sozialmedizinische Empfehlungen aufgrund einer Grundsatzbegutachtung nach EbM-Kriterien zeigen begründet auf, ob und ggf. wie neue Methoden sachgerecht in die Versorgung eingeführt werden können.

Literatur

- [1] Mayer ED et al., Die evidenzbasierte wissenschaftliche Analyse als Grundlage MDK-interner Grundsatzgutachten. Z Arztl Fortbild Qualitätssich. (2001), 95, 113–119
- [2] Medizinischer Dienst der Spitzenverbände der Krankenkassen e.V. (MDS). Begutachtungsanleitung „Neue Untersuchungs- und Behandlungsmethoden (NUB)“ ohne Fertigarzneimittel. http://www.mds-ev.org/download/BGA-NUB_2002.pdf (01.12.06)
- [3] Medizinischer Dienst der Spitzenverbände der Krankenkassen e.V. (MDS). Empfehlungen zur Erstellung gutachterlicher Stellungnahmen zur Bewertung medizinischer Untersuchungs- und Behandlungsmethoden. [http://infomed.mds-ev.de/sindbad.nsf/2d4c1f1235206c83c12571e700442be1/a58e2af4d44f7cd300256c6100548108/\\$FILE/P41_Standards.pdf](http://infomed.mds-ev.de/sindbad.nsf/2d4c1f1235206c83c12571e700442be1/a58e2af4d44f7cd300256c6100548108/$FILE/P41_Standards.pdf) (01.12.06)
- [4] Giehl J, Methodenbewertung Vertebroplastie. Kurzfassung. <http://www.mds-ev.org/veranstaltungen/abstracts/16022006/Abstract5-Giehl.pdf> (01.12.06)
- [5] Felsenberg D et al., Prävalenz der vertebrealen Wirbelkörperdeformationen bei Frauen und Männern in Deutschland. EVOS-Gruppe in Deutschland. Med Klin (Munich). (1998), 93 Suppl 2, 31–34
- [6] National Institute for Clinical Excellence (NICE). Percutaneous vertebroplasty. <http://www.nice.org.uk/guideline/IPG12> (01.12.06)
- [7] Diamond TH, Champion B, Clark WA, Management of acute osteoporotic vertebral fractures: a nonrandomized trial comparing percutaneous vertebroplasty with conservative therapy. Am J Med. (2003), 114, 257–265
- [8] Kasperk C et al., Treatment of painful vertebral fractures by kyphoplasty in patients with primary osteoporosis: a prospective nonrandomized controlled study. J Bone Miner Res. (2005), 20, 604–612
- [9] National Institute for Clinical Excellence (NICE). Balloon kyphoplasty for vertebral compression fractures. <http://www.nice.org.uk/guidance/IPG166> (01.12.06)

34 Evidenzbasierte Medizin in der Zahnmedizin – Beispiel Myoarthropathien des Kausystems

Jens C. Türp

„Wo ist der Beweis?“

(Wilhelm Wallisch,

Privatdozent für Zahnheilkunde an der Universität Wien, im Jahre 1911 in der Österreichischen Zeitschrift für Stomatologie)

Die Diagnostik und Therapie der Myoarthropathien des Kausystems (MAP) gehören zu jenen Gebieten innerhalb der klinischen Zahnmedizin, die mit besonders vielen Kontroversen verbunden sind. Der Terminus „Myoarthropathien“ steht als Sammelbegriff für eine Reihe von (vornehmlich) schmerzbezogenen Problemen im Bereich der Kau-muskulatur, Kiefergelenke und benachbarten Strukturen. Myoarthropathien werden als eine Form des muskuloskelettalen Schmerzes angesehen [1]. Neben Schmerzen, von denen der Patient im Rahmen der Krankengeschichte sowie als Folge der Palpation von Kaumuskeln und Kiefergelenken berichtet, können Einschränkungen der Unterkieferbeweglichkeit vorhanden sein; diese sind oftmals eine Folge der Schmerzen (Schonhaltung).

Speziell in Deutschland wird vielfach der Begriff „craniomandibuläre Dysfunktionen“ (abgekürzt: CMD) verwendet. Dieser ist nicht unumstritten, da weder der Terminus „Dysfunktion“ noch diagnostik- und behandlungsbedürftige Inhalte desselben klar definiert sind.

34.1 Das Versorgungsproblem

MAP stellen ein ernst zu nehmendes Schmerzproblem dar; die Prävalenz myoar-

thropathischer Schmerzen in der erwachsenen Bevölkerung liegt in Deutschland bei knapp 5%, wobei Frauen häufiger betroffen sind als Männer.

Diagnostik und Behandlung von MAP-Patienten sind durch eine erstaunliche Praxisvariation gekennzeichnet. Das in der deutschsprachigen Zahnmedizin langsam erwachende Interesse an den Prinzipien der Evidenzbasierten Medizin [2] hat deutlich werden lassen, dass bis heute diagnostische und therapeutische Methoden zum Einsatz kommen, für deren Wirksamkeit keine überzeugenden Belege vorhanden sind. Dies kann zu Über-, Unter- und Fehlversorgung führen [3; 4].

34.2 Klinisches Szenario

Frau Zuberbühler, 35 Jahre alt, alleinerziehende Mutter eines Kleinkinds, empfindet seit einem halben Jahr ein Spannungsgefühl und dumpfe, drückende Schmerzen in beiden Masseteren und den Schläfenmuskeln (Mm. temporales). Die Schmerzen sind nach dem morgendlichen Erwachen kaum vorhanden, nehmen aber im Laufe des Tages zu. Kauen insbesondere von harten und zähen Speisen führt zu einer Verstärkung der Beschwerden. Die Schmerzsymptomatik ist für einige Wochen spontan geringer geworden, hat seit einem Monat aber wieder zugenommen. Da Frau Zuberbühler darüber hinaus eine Veränderung im Zusammenbiss ihrer Zähne bemerkt hat („Wenn ich Schmerzen habe, passen die Zähne oft nicht mehr so gleichmäßig aufeinander wie vorher“), sucht sie ihren Hauszahnarzt auf. Dieser erklärt ihr nach einer zweiminütigen extra- und intra-



Abb. 34.1: Attrition (Zahnschmelzverlust) aller Frontzähne als Folge langjährigen Zähneknirschens. Die ehemaligen Eckzahnspitzen haben sich dadurch in Schneidekanten umgewandelt. Man beachte – hier bei nach links geschobenem Unterkiefer – die spaltfreie Passung der Schneidekanten der Frontzähne.

ralen Untersuchung, die Schmerzen seien durch Verspannungen der Kaumuskulatur zustande gekommen. Ursache der Muskelverspannung sei zum einen Disstress, zum anderen eine fehlerhafte Okklusion der Zähne mit Frühkontakten beim Zusammenbeißen, was zu nächtlichem Zähneknirschen führe; in der Tat sind die abgeschliffenen Frontzähne deutlich zu erkennen (s. Abb. 34.1).

Der Zahnarzt gliedert im Oberkiefer eine so genannte Michigan-Schiene (s. Abb. 34.2, 34.3) ein. Frau Zuberbühler soll diese Kunststoffschiene nachts tragen.

Zur Unterstützung der Therapie verschreibt ihr der Zahnarzt (die in der Schweiz besonders beliebte) Mefenaminsäure (Ponstan, 500 mg, 1-1-1). Der Zahnarzt schlägt ferner vor, nach Abklingen der Schmerzen die Zähne so einzuschleifen, dass beim Kieferschluss und bei zahngeführten Unterkieferbewegungen nach vorne und zur Seite auf allen Zähnen gleichmäßige Kontakte erzielt werden. Diese Therapiemaßnahme, mit der er seit über 20 Jahren gute persönliche Erfah-

rungen gemacht habe, verhindere ähnliche Beschwerden in der Zukunft. Frau Zuberbühler reagiert auf diesen Behandlungsvorschlag skeptisch und lehnt ihn ab.

Die Mefenaminsäure nimmt sie wie verordnet ein, setzt sie aber nach drei Wochen wegen auftretender Magenschmerzen und leichter Übelkeit ab. Da sich die Schmerzsymptomatik auch vier Wochen nach Behandlungsbeginn kaum gebessert hat, bittet Frau Zuberbühler um einen Termin in der Klinik für Rekonstruktive Zahnmedizin und Myoarthropathien der Universitätsklinik für Zahnmedizin in Basel.

34.3 Die Evidenz bewerten

Zu den typischen Maßnahmen, die von Zahnärzten zur Behandlung von Kaumuskelschmerzen (myofaszialen Schmerzen) durchgeführt oder verordnet werden, zählen:

- ▲ Aufklärung/Verhaltensmaßnahmen,
- ▲ Schienentherapie,

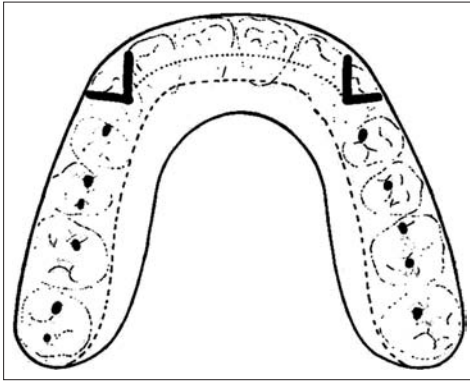


Abb. 34.2: Schematische Darstellung einer Michigan-Schiene in der Aufsicht: Die Punkte stellen die beim Zubeißen auftretenden Kontakte der unteren Seiten- und Eckzähne auf der Schiene dar. Bei Vor- und Seitschub des Unterkiefers entlang der Schienenoberfläche haben nur die unteren Eckzähne Kontakt (V-förmige Markierung).

- ▲ Krankengymnastik/Physiotherapie,
- ▲ Akupunktur,
- ▲ Pharmakologische Therapie,
- ▲ Psychologische Therapie,
- ▲ Systematisches Einschleifen des Gebisses.

Abhängig von der klinischen Situation (und der persönlichen Vorliebe des Zahnarztes) werden diese Therapieoptionen einzeln oder in Kombination angewandt. Im Folgenden soll für die Schienenbehandlung beispielhaft gezeigt werden, ob Evidenz vorliegt, die diese Therapiemaßnahme rechtfertigt.

34.3.1 Schientherapie

Unter den vielfältigen Schientypen, die in den vergangenen 120 Jahren beschrieben worden sind, wird die so genannte Michigan-Schiene von vielen Klinikern seit langem für die Behandlung von Myalgien der Kaumuskeln, Arthralgien der Kiefergelenke sowie von Bruxismus (Zähneknirschen; Kieferpressen) empfohlen und eingesetzt. Die aus hartem Kunststoff bestehende Okklusionsschiene – sie wurde in den 1960er-Jahren an der University of Michigan (Ann Arbor) entwickelt –



Abb. 34.3: „Prinzip der Michigan-Schiene: Eine Erhöhung des Kunststoffs im Eckzahnbereich sorgt dafür, dass bei Vor- und Seitschubbewegungen des Unterkiefers ausschließlich die Eckzähne Kontakt mit der Schienenoberfläche haben (Eckzahnführung); im Bereich der Schneide- und Seitenzähne kommt es zu einem Klaffen (Disklusion). Hier abgebildet: Eckzahnführung bei Unterkiefer-Seitschub nach links.

bedeckt die Kauflächen und einen Teil der Bukkal- bzw. Labialflächen aller Oberkieferzähne, ferner den vorderen Gaumenbereich (vgl. Abb. 34.2). Für die Herstellung einer Schiene macht der Zahnarzt zunächst eine Abformung der Zahnreihen des Ober- und Unterkiefers mit Alginat. Anschließend stellt der Zahntechniker im Zahnlabor Ober- und Unterkiefermodelle aus Gips her, modelliert die Schiene in Wachs und ersetzt dieses dann durch (harten) Kunststoff. Im Gegensatz zu manch anderen Schienenarten ist bei Anwendung der Michigan-Schiene keine Gefahr ernsthafter Nebenwirkungen (z.B. irreversible Veränderungen der Okklusion der Zähne) zu befürchten.

Bei Patienten mit myofaszialen Schmerzen der Kaumuskulatur ist das regelmäßige Tragen einer Michigan-Schiene in der Regel mit einem Schmerzrückgang (nicht gleichbedeutend mit Schmerzfreiheit) verbunden [5]. Dabei scheinen Patienten, deren Schmerzen auf die Kiefermuskulatur beschränkt bleiben, von der Schientherapie mehr zu profitieren als Patienten mit zusätzlichen Schmerzen in anderen Körperbereichen [6]. Es ist nicht

abschließend geklärt, ob die zu beobachtende Schmerzlinderung eher durch unspezifische Effekte [7] oder durch eine spezifische Schienenwirkung [6; 8] zustande kommt.

Wichtig ist die Beobachtung, dass einige andere Behandlungsmaßnahmen zu ähnlichen Ergebnissen führen wie die Therapie mit einer Michigan-Schiene; in randomisierten kontrollierten Studien wurde mit folgenden Therapiemitteln ein vergleichbar großer Schmerzurückgang erreicht [5]:

- ▲ weiche Kunststoffschiene [9],
- ▲ nur den Gaumen bedeckende (harte) Schiene [10],
- ▲ Physiotherapie,
- ▲ Körperakupunktur,
- ▲ systematisches okklusales Einschleifen der Zähne.

Klinisches Szenario

Die fehlende Überlegenheit der Michigan-Schiene gegenüber anderen Behandlungsmaßnahmen spricht nicht gegen ihren Einsatz bei myofaszialen Schmerzen der Kaumuskulatur. Für die (praktisch nebenwirkungsfreie) Michigan-Schiene lassen sich – neben ihrer mehr als 40-jährigen weltweiten Bewährung – der im Vergleich zu Physiotherapie- und Akupunktur-Sitzungen geringere zeitliche Aufwand für den Patienten ins Feld führen. Ein weiterer Vorteil ist ihre Nicht-Invasivität.

34.4 Diskussion

Aufgrund des weiterhin bestehenden Mangels an qualitativ hochwertigen klinischen Studien ist die praktische Durchführung einer auf hoher Evidenz stehenden Versorgung von MAP-Patienten eingeschränkt. Daher muss die Anwendung der meisten Behandlungsmaßnahmen, die derzeit zur Therapie akuter und persistierender Kaumuskel- und/oder Kiefergelenkschmerzen zum Einsatz kommen, als „nicht validierte klinische Praxis“ angesehen werden [11]. Dies gilt auch für einen großen Teil der für die Therapie

muskuloskelettaler Gesichtsschmerzen eingesetzten Arzneimittel [12].

In dieser Situation scheinen viele Zahnärzte die auf persönlicher klinischer Erfahrung beruhenden Aussagen respektierter Autoritäten als gleichwertigen Ersatz anzusehen. Da diese Art der Wissensübermittlung und -aneignung derjenigen entspricht, die traditionell während des Studiums an der Universität erfolgte („ex cathedra“), wird vielfach keine Veranlassung gesehen, sich kritisch mit der Qualität der erhaltenen Informationen auseinanderzusetzen – der Respekt vor dem guten Ruf des Experten genügt als Qualitäts- und Wahrheitsbeweis.

Ein solches blindes Vertrauen erklärt, warum beispielsweise die Bemerkung, nach systematischer Einschleiftherapie seien myoarthropathische Schmerzen „almost always resolvable in a matter of minutes once the occlusion has been refined as well as possible“ [13], über Jahrzehnte kaum angezweifelt wurde. Aus den Ergebnissen der vorhandenen randomisierten kontrollierten klinischen Studien und umfassenden bevölkerungsrepräsentativen epidemiologischen Untersuchungen [14] ergeben sich jedenfalls **keine** überzeugenden Hinweise für die von vielen Zahnärzten seit mehr als einem halben Jahrhundert gehegten Vermutungen, dass (1) okklusale Faktoren (z.B. störende Vorkontakte beim Zubeißen) ein Hauptgrund für die Entstehung und Aufrechterhaltung von Schmerzen im Bereich der Kaumuskeln oder Kiefergelenke seien und (2) ein mit dem Ziel einer „**okklusalen Harmonisierung und Stabilisierung**“ [15] durchgeführtes systematisches Einschleifen des Gebisses mit einem klinischen Nutzen verbunden sei. Daher wird dieses systematische Beschleifen der Zähne weder für die Behandlung noch für die Prävention von MAP empfohlen [16]. Auch die verbreitete Annahme, okklusale Störkontakte seien die Ursache für Bruxismus, kann in Kenntnis der Ergebnisse einer vor kurzem veröffentlichten randomisierten,

placebokontrollierten Doppelblindstudie nicht länger aufrechterhalten werden [17].

Demgegenüber haben experimentelle Studien gezeigt, dass die von einigen MAP-Patienten angegebenen und bei einer intraoralen Untersuchung auch tatsächlich feststellbaren okklusalen Dysharmonien **Folge** der Schmerzen sein können: Infusion einer algetischen Substanz (5%ige NaCl-Lösung) in den Musculus masseter führt zu einer schmerzbedingten Lageveränderung des Unterkiefers (Schonhaltung) und damit der unteren Zähne, was die wahrgenommenen okklusalen Diskrepanzen beim Zusammenbeißen erklärt [18]. Eine Einschleiftherapie wäre in dieser Situation ein zahnärztlicher Kunstfehler.

Ein eventuelles Ausbleiben von Behandlungserfolgen nach systematischem Einschleifen wird von Vertretern der traditionellen Ansichten in der Regel nicht der fehlenden Wirkung der Behandlungsmaßnahme angelastet, sondern der mangelnden Geschicklichkeit des Zahnarztes: „Dentists, whose patients are not routinely experiencing this dramatic remission of pain after occlusal equilibration, are simply not equilibrating them properly“ [13]. Eine ähnliche Argumentation wird für die Erfolgswahrscheinlichkeit einer Schienentherapie gewählt: „Like everything else in dentistry; success is related to operator skill“ [19]. Zweifellos stimmt solch eine Feststellung für eine Zahnpräparation, eine gelegte Füllung, eine zahnprothetische Versorgung, ein gesetztes Implantat oder einen oralchirurgischen Eingriff – nicht aber, wenn es um die Behandlung von Patienten mit persistierenden Schmerzen geht. Ein berichteter/beobachteter Therapieerfolg, wie Schmerzurückgang oder verbesserte Kieferfunktion, kann durchaus ganz andere Ursachen haben [20], wie

- ▲ spontane Beschwerdeverbesserung,
- ▲ das statistische Phänomen des Rückgangs (Regression) zum Mittelwert,
- ▲ günstige, auf die Person des Behandlers, die Art der angewandten Prozedur oder

die Umgebung zurückzuführende Einflüsse (Placebo-Effekt),

- ▲ die Erwartung des Patienten, durch die Behandlung einen Rückgang der Beschwerden zu erfahren,
- ▲ Verhaltensänderungen des Patienten aufgrund des Hawthorne-Effekts,
- ▲ (unbeachtete) begleitende Therapiemaßnahmen,
- ▲ gefällige Antworten des Patienten.

Hat sich der behandelnde Zahnarzt aber einmal einem bestimmten Therapiekonzept verschrieben, so ist er anfällig für den Trugschluss „post hoc ergo propter hoc“: Er wird geneigt sein, den beobachteten Schmerzurückgang allein seiner Behandlungsstrategie zuzuschreiben. Spezifische Therapiewirkungen lassen sich aber nur mithilfe von randomisierten kontrollierten Studien nachweisen.

Eine besonders kritische Haltung sollte man – auch bei nachgewiesener hoher Evidenz – gegenüber irreversiblen Behandlungsmethoden einnehmen: es müssen Belege dafür vorhanden sein, dass der Nutzen der Intervention (a) größer ist als derjenige, der mit einer nicht-invasiven Vorgehensweise oder ohne Therapie erzielt wird, **und** (b) in einem sinnvollen Verhältnis steht zu dem Aufwand, den Risiken und den Kosten (sowie den eventuell entstehenden Folgekosten). Solche Nachweise liegen zum Beispiel für die systematische Einschleiftherapie nicht vor [16].

34.5 Auflösung des klinischen Szenarios

Zunächst wurden standardisiert die Krankengeschichte erhoben und eine klinische Untersuchung durchgeführt. Mit Hilfe von validierten Fragebögen wurde die Auswirkung der persistierenden Muskelschmerzen auf das seelische Wohlbefinden und auf die Verrichtung täglicher Aktivitäten erfragt. Da Frau Zuberbühler angab,

dass sie manchmal auch in anderen Körperbereichen Muskel- und Gelenkschmerzen spürt, wurde mit ihrem Hausarzt Rücksprache gehalten, um zu erfahren, ob weitere Schmerzdiagnosen bekannt sind (z.B. Fibromyalgie-Syndrom), was dieser aber verneinte. Bei dieser Gelegenheit wurde dem Hausarzt das geplante weitere Vorgehen mitgeteilt.

Angesichts der großen Bedeutung einer genauen Aufklärung für den Therapieerfolg [21] wurden Frau Zuberbühler detailliert und in verständlichen Worten die Befunde und die Diagnose („myofaszialer Schmerz der Kiefermuskulatur“), die mutmaßlichen Ursachen und die zur Verfügung stehenden Behandlungsmöglichkeiten erläutert. Ihr wurde erklärt, dass die klinische Symptomatik – kaum Schmerzen am Morgen, Zunahme der Beschwerden mit fortschreitendem Tag – nicht darauf hindeutet, dass nächtliches Zähneknirschen oder Kieferpressen die Ursache ihrer Beschwerden sei. Ferner wurde Frau Zuberbühler über den derzeitigen Wissensstand zur Rolle der Okklusion bei Kaumuskelschmerzen informiert. Deutlich wurde auf die Risiken hingewiesen, die mit der Anwendung irreversibler Therapien verbunden sind. Sie wurde auch darüber in Kenntnis gesetzt, dass keine Evidenz aus qualitativ hoch stehenden Studien vorliegt, die die Wirksamkeit des Anthranilsäurederivats Mefenaminsäure oder anderer Nichtopioid-Analgetika bei myofaszialen Schmerzen der Kaumuskulatur belegt, und dass saure Analgetika (nicht-steroidale Antirheumatika) nur zur Therapie akuter Kiefergelenkschmerzen empfohlen werden [11].

Auf der Okklusalfäche der von Frau Zuberbühlers Hauszahnarzt hergestellten Michigan-Schiene wurden größere Schleifkorrekturen vorgenommen, um sicherzustellen, dass die unteren Eck- und Seitenzähne beim gewohnten Zubeißen gleichmäßig und gleichzeitig mit der Oberkieferschiene in Kontakt kommen. Um einen durch nächtliches Zähneknirschen bedingten Verlust von Zahnhartsubstanz zu verhindern, wurde der Patientin geraten, die Schiene weiterhin nachts zu tragen, im Falle stärkerer Stress-Episoden ggf. zusätzlich einige Stunden am Tage.

Da sich in mehreren randomisierten klinischen Studien zeigte, dass MAP-Patienten durch eine zusätzliche psychologische Schmerztherapie (z.B. Entspannungstraining und Stressmanagement; kognitive Verhaltenstherapie) einen größeren Behandlungserfolg erwarten können als bei ausschließlich somatischer Therapie [22; 23], wurde Frau Zuberbühler die Adresse eines Schmerzpsychotherapeuten gegeben.

Eine mögliche Verordnung von niedrig dosiertem Amitriptylin, das – wie andere trizyklische Antidepressiva – wegen seiner direkten und indirekten analgetischen Wirkung zur länger dauernden Behandlung persistierender nicht-maligner Schmerzen angewandt wird, wurde von den Ergebnissen der eingeleiteten Behandlungsmaßnahmen abhängig gemacht.

34.6 Schlussfolgerungen

In der ersten Auflage dieses Lehrbuchs wurde ein vergleichbares klinisches Szenario vorgestellt wie in diesem Kapitel [24]. In dem seit her vergangenen Zeitraum sind – vom Gros der Zahnärzte weitgehend unbeachtet – wichtige systematische Übersichten und methodisch ausgezeichnete Artikel über diagnostische und therapeutische Studien zum Thema MAP veröffentlicht worden. Dadurch ist das Fundament, auf dem patientenbezogene Entscheidungen getroffen werden können, verbreitert und qualitativ verbessert worden – auch wenn weiterhin große Lücken bestehen.

Die erzielten Fortschritte nützen allerdings wenig, wenn es nicht gelingt, mehr als nur die besonders motivierten Kollegen von den Vorteilen einer Evidenzbasierten (Zahn-)Medizin zu überzeugen. Trotz der zunehmenden Zahl publizierter Beiträge zum Thema Evidenzbasierte Zahnmedizin [2] bleiben viele Zahnärzte zurückhaltend [25]. Es fällt ihnen anscheinend schwer, anzuerkennen, dass persönliche Eindrücke und „charisambasierte Behandlungsphilosophien“ [11] merkbar an Bedeutung verloren haben und dass sich „Wis-

Tab. 34.1: Informationsangebote zum Thema Evidenzbasierte Zahnmedizin

Quelle	Netzseite
Centre for Evidence-Based Dentistry, Universität Oxford	www.cebd.org
DSM-Forsyth Center for Evidence-Based Dentistry, The Forsyth Institute, Boston, Massachusetts	www.forsyth.org/forsyth.asp?pg=100111
American Dental Association	www.ada.org/prof/resources/positions/statements/evidencebased.asp
New Zealand Dental Association	www.nzda.org.nz/evidence/ebd_home.htm
International Centre for Evidence-Based Oral Health (ICEBOH), Eastman Dental Institute, London	www.eastman.ucl.ac.uk/iceboh

sen“, das früher in so genannten „Schulen“ ohne methodisch adäquate Überprüfung weitergegeben wurde, heute in hochqualitativen klinischen Studien beweisen muss.

Daher müssen auch Lösungen gefunden werden, wie den zahnärztlichen Kolleginnen und Kollegen aktuelle Studienergebnisse zeitnah zur Kenntnis gebracht werden können. Geeignete Quellen sind vorhanden: **Eine** empfehlenswerte Möglichkeit, das persönliche Fachwissen auf den aktuellen Stand zu bringen und dort zu halten, ist die regelmäßige Lektüre der Zeitschriften „Evidence-Based Dentistry“ und „The Journal of Evidence Based Dental Practice“. Zur Aneignung des Handwerkszeugs der Evidenzbasierten (Zahn-) Medizin dienen unter anderem die „EbM-Splitter“, die seit Februar 2001 zweimonatlich in der „Deutschen Zahnärztlichen Zeitschrift“ erscheinen. Auch Bücher [26–28] und Buchkapitel [29] über das Thema Evidenzbasierte Zahnmedizin liegen vor. Empfehlenswerte Informationsquellen aus dem Internet sind in Tabelle 34.1 zusammengefasst. Es besteht eigentlich kein rationaler Grund für eine skeptische Haltung gegenüber einer Evidenzbasierten Zahnmedizin. Oder doch?

*„Eigentlich weiß man nur,
wenn man wenig weiß;
mit dem Wissen wächst der Zweifel.“*
(Johann Wolfgang von Goethe,
Maximen und Reflexionen,
1833 [postum veröffentlicht])

Literatur

- [1] Svensson P (2002) Orofacial musculoskeletal pain. In: Giamberardino MA, Pain 2002 – An updated review, 447–458. IASP Pr., Seattle
- [2] Türp JC, Röhrich C, Antes G, Evidence-based dentistry in German-language dental journals 1997–2003. Evid Based Dent. (2005), 6, 86–87
- [3] Türp JC, Über-, Unter- und Fehlversorgung in der Funktionsdiagnostik und -therapie. Beispiele, Gefahren, Gründe. Teil I. Schweiz Monatsschr Zahnmed. (2002), 112, 819–829
- [4] Türp JC, Über-, Unter- und Fehlversorgung in der Funktionsdiagnostik und -therapie. Beispiele, Gefahren, Gründe. Teil II. Schweiz Monatsschr Zahnmed. (2002), 112, 908–919
- [5] Türp JC, Komine F, Hugger A, Efficacy of stabilization splints for the management of patients with masticatory muscle pain: a qualitative systematic review. Clin Oral Investig. (2004), 8, 179–195
- [6] Raphael KG, Marbach JJ, Widespread pain and the effectiveness of oral splints in myofascial face pain. J Am Dent Assoc. (2001), 132, 305–316
- [7] Dao TT et al., The efficacy of oral splints in the treatment of myofascial pain of the jaw muscles: a controlled clinical trial. Pain. (1994), 56, 85–94
- [8] Ekberg E, Vallon D, Nilner M, The efficacy of appliance therapy in patients with temporomandibular disorders of mainly myogenous origin. A randomized, controlled, short-term trial. J Orofac Pain. (2003), 17, 133–139
- [9] Truelove E et al., The efficacy of traditional, low-cost and nonsplint therapies for

- temporomandibular disorder: a randomized controlled trial. *J Am Dent Assoc.* (2006), 137, 1099–1107
- [10] Wassell RW, Adams N, Kelly PJ, The treatment of temporomandibular disorders with stabilizing splints in general dental practice: one-year follow-up. *J Am Dent Assoc.* (2006), 137, 1089–1098
- [11] Dionne RA (2006) Pharmacologic approaches. In: Laskin DM, Greene CS, Hylander WL, TMDs. An Evidence-based approach to diagnosis and treatment, 347–357. Quintessence, Chicago
- [12] List T, Axelsson S, Leijon G, Pharmacologic interventions in the treatment of temporomandibular disorders, atypical facial pain, and burning mouth syndrome. A qualitative systematic review. *J Orofac Pain.* (2003), 17, 301–310
- [13] Dawson PE, Temporomandibular joint pain-dysfunction problems can be solved. *J Prosthet Dent.* (1973), 29, 100–112
- [14] Gesch D et al., Association of malocclusion and functional occlusion with subjective symptoms of TMD in adults: results of the Study of Health in Pomerania (SHIP). *Angle Orthod.* (2005), 75, 183–190
- [15] Engelhardt JP (1995) Einschleiftherapie im natürlichen Gebiß. In: Koeck B, Funktionsstörungen des Kauorgans, 243–279. Urban und Schwarzenberg, München
- [16] Koh H, Robinson PG, Occlusal adjustment for treating and preventing temporomandibular joint disorders. *Cochrane Database Syst Rev.* (2003), CD003812
- [17] Michelotti A et al., Effect of occlusal interference on habitual activity of human masseter. *J Dent Res.* (2005), 84, 644–648
- [18] Obrez A, Stohler CS, Jaw muscle pain and its effect on gothic arch tracings. *J Prosthet Dent.* (1996), 75, 393–398
- [19] Ramfjord SP, Ash MM, Reflections on the Michigan occlusal splint. *J Oral Rehabil.* (1994), 21, 491–500
- [20] Türp JC, Schwarzer G, Zur Wirksamkeit therapeutischer Massnahmen: der Post-hoc-ergo-propter-hoc-Trugschluss. *Schweiz Monatsschr Zahnmed.* (2003), 113, 36–46
- [21] Michelotti A et al., The additional value of a home physical therapy regimen versus patient education only for the treatment of myofascial pain of the jaw muscles: short-term results of a randomized clinical trial. *J Orofac Pain.* (2004), 18, 114–125
- [22] Dworkin SF et al., A randomized clinical trial using research diagnostic criteria for temporomandibular disorders-axis II to target clinic cases for a tailored self-care TMD treatment program. *J Orofac Pain.* (2002), 16, 48–63
- [23] Turk DC et al., Dysfunctional patients with temporomandibular disorders: evaluating the efficacy of a tailored treatment protocol. *J Consult Clin Psychol.* (1996), 64, 139–146
- [24] Türp JC, Antes G (2000) EbM in der Zahnmedizin-Beispiel „Myoarthropathien des Kausystems“. In: Kunz R, Ollenschläger G, Raspe HH, Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis, 330–338. Dt. Ärzte-Verl., Köln
- [25] Donner-Banzhoff N, Evidenzbasierte (Zahn-)Medizin-eine apologetische Betrachtung. *Dtsch Zahnärztl Z.* (2006), 61, 321–323
- [26] Böning K, Kirch W (Ed.) (2003) Evidenzbasierte Zahnheilkunde. Zahngesundheit, Patientenorientierung, Leitlinien. Quintessenz-Verl., Berlin
- [27] Clarkson J et al. (2003) Evidence based dentistry for effective practice. Martin Dunitz, London
- [28] Walther W, Micheelis W (Ed.) (2000) Evidence-Based Dentistry. Evidenz-basierte Medizin in der Zahn-, Mund- und Kieferheilkunde. Dt. Zahnärzte-Verl., Köln
- [29] Türp JC, Böning KW (2002) Evidenzbasierte Zahnmedizin. In: Perleth M, Antes G, Evidenz-basierte Medizin. Wissenschaft im Praxisalltag, 82–90. Urban und Vogel, München

IV EbM in den Strukturen der Gesundheitsversorgung

35	Evidenzbasierte Medizin und Leitlinien	361
	<i>Ina Kopp, Monika Lelgemann, Günter Ollenschläger</i>	
36	Systemsteuerung im Rahmen des SGB V: der Gemeinsame Bundesausschuss	375
	<i>Bernhard Gibis, Regina Klakow-Franck, Nicole Schlottmann, Johannes Bruns</i>	
37	Evidenzbasierte Medizin und ethische Aspekte	387
	<i>Nikola Biller-Andorno, Christian Lenk</i>	
38a	Evidenzbasierte Medizin und Recht: Haftungsrecht	393
	<i>Dieter Hart</i>	
38b	Evidenzbasierte Medizin und Recht: Sozialrecht	401
	<i>Felix Welti</i>	

35 Evidenzbasierte Medizin und Leitlinien

Ina Kopp, Monika Lelgemann, Günter Ollenschläger

In zahlreichen Kapiteln des vorliegenden Buchs wird auf medizinische Leitlinien Bezug genommen. Dies entspricht ihrer Bedeutung im Rahmen der evidenzbasierten Gesundheitsversorgung und des Qualitätsmanagements [1]. Der Transfer relevanter Informationen aus der Wissenschaft in den Versorgungsalltag stellt sich angesichts der raschen Entwicklung neuer medizinischer Technologien und der Überflutung mit wissenschaftlichen Publikationen immer schwieriger dar. Leitlinien können diesem Dilemma begegnen, indem sie „die beste Evidenz mit anderen Kenntnissen, die für die Entscheidungsfindung hinsichtlich eines bestimmten Gesundheitsproblems erforderlich sind“, [2] kombinieren und als Vektoren für die Umsetzung dieses besten verfügbaren Wissens genutzt werden.

Mittlerweile sind Leitlinien nicht mehr aus dem medizinischen Alltag wegzudenken und werden auch in Zukunft das Handeln im Gesundheitswesen maßgeblich beeinflussen. Im folgenden Kapitel wird eine Übersicht über die Ziele, die Entwicklung und Nutzung von Leitlinien gegeben.

35.1 Was sind Leitlinien und wozu dienen sie?

Vorrangiges Ziel von Leitlinien ist die Verbesserung der Qualität medizinischer Versorgung durch Wissensvermittlung. Leitlinien sollen auf der Basis des besten verfügbaren Wissens definieren, was optimale Versorgungsqualität bedeutet, unter Abwägung von Nutzen und Schaden alternativer Vorge-

hensweisen und unter Berücksichtigung der Bedürfnisse und Einstellungen der Patienten (zu den Definitionen s. Tab. 35.1) [3–8]. Leitlinien sind weder als Anleitung für eine „Kochbuchmedizin“ zu verstehen, noch stellen sie die Meinungen einzelner Fachexperten dar [9]. Leitlinien sollen als systematisch entwickelte Entscheidungshilfen konkrete Handlungsempfehlungen für den Versorgungsablauf bereitstellen. Dadurch unterscheiden sie sich von Evidenzberichten wie Systematischen Übersichtsarbeiten und Health Technology Assessment [10; 11]. Leitlinien sind im Gegensatz zu Richtlinien nicht verbindlich [12]. Die Anwendbarkeit einer bestimmten Empfehlung in der individuellen Situation ist unter Berücksichtigung der vorliegenden Gegebenheiten (z.B. Begleiterkrankungen des Patienten, verfügbare Ressourcen) zu prüfen .

Aus Leitlinien können Zielgrößen und Indikatoren abgeleitet werden, anhand derer die Qualität der medizinischen Versorgung messbar und eine Unterscheidung zwischen gut und verbesserungsbedürftig ermöglicht wird. Solche Qualitätsindikatoren sind ein wesentliches Werkzeug für die Evaluierung von Versorgungsleistungen und -ergebnissen im medizinischen Alltag, für das interne Qualitätsmanagement und für den externen Qualitätsvergleich [13].

In der „Qualitätsmanagement-Richtlinie für die vertragsärztliche Versorgung“ hat auch der Gemeinsame Bundesausschuss (G-BA) den Stellenwert von Leitlinien als einem Grundelement des einrichtungsinternen Qualitätsmanagements betont [14].

Tab. 35.1: Leitlinien – Definitionen (nach [3–8])**Leitlinien**

- sind definiert als „systematisch entwickelte Aussagen zur Unterstützung der Entscheidungsfindung von Ärzten und ggf. anderen Gesundheitsberufen sowie Patienten für eine angemessene Vorgehensweise bei speziellen Gesundheitsproblemen“;
- sind Orientierungshilfen im Sinne von „Handlungs- und Entscheidungskorridoren“, von denen in begründeten Fällen abgewichen werden kann oder sogar muss;
- haben die Aufgabe, das umfangreiche Wissen (wissenschaftliche Evidenz und Praxiserfahrung) zu speziellen Versorgungsproblemen
 - explizit darzulegen,
 - unter methodischen und klinischen Aspekten zu bewerten,
 - gegensätzliche Standpunkte zu klären sowie
 - unter Abwägung von Nutzen und Schaden das derzeitige Vorgehen der Wahl zu definieren;
- zielen darauf, unter Berücksichtigung der vorhandenen Ressourcen
 - gute klinische Praxis zu fördern und die Öffentlichkeit darüber zu informieren,
 - Entscheidungen in der medizinischen Versorgung auf eine rationalere Basis zu stellen,
 - die Stellung des Patienten als Partner im Entscheidungsprozess zu stärken und
 - die Qualität der Versorgung zu verbessern.

Schließlich können Leitlinien über die Identifizierung und Abschaffung unwirksamer und veralteter Verfahren zu einer angemesseneren Verwendung finanzieller Ressourcen beitragen. Die aktive Nutzung von Leitlinien kann jedoch auch dazu führen, dass spezielle Interventionen und Therapien auf der Grundlage besonders zuverlässiger Evidenz vermehrt angeboten und in Anspruch genommen werden.

Der günstige Einfluss von Leitlinien auf die Prozess- und Ergebnisqualität im Gesundheitswesen ist mittlerweile wissenschaftlich belegt [15; 16]. Die Wirksamkeit und damit letztlich der Nutzen einer einzelnen Leitlinie hängt aber entscheidend von ihrer Qualität und von ihrer Umsetzung ab [17; 18]. Demnach werden heute international bestimmte Kriterien, die hochwertige Leitlinien erfüllen sollten, in einheitlicher Weise definiert [1; 3–6; 19]. Für den deutschen Raum sind diese Kriterien methodischer Qualität in Form einer kommentierten Checkliste publiziert, dem Deutschen Leitlinienbewertungs-Instrument (DELBI) [20] (s. Kap. 15). DELBI soll einerseits Leitlinienanwendern und anderen Interessierten helfen,

Leitlinien einer methodischen Überprüfung zu unterziehen. Andererseits soll es aber auch Leitlinienentwickler unterstützen – zum einen bei der Selbsteinschätzung, zum anderen, um sicherzustellen, dass ihre Leitlinien den internationalen Standards entsprechen.

Dabei sind zunächst drei grundlegende Aspekte hervorzuheben [21; 22]:

- ▲ Zusammensetzung des Leitliniengremiums: Repräsentativität für den Anwenderkreis,
- ▲ Evidenzbasierung: systematische Suche, Auswahl und Bewertung der Literatur,
- ▲ Methodik der Entwicklung: systematische Evidenz- und Konsensbasierung.

Die Evidenzbasierung ist vor allem maßgeblich für die wissenschaftliche Legitimation einer Leitlinie, während die Beteiligung der Anwender sowie die strukturierte Konsensfindung vor allem für die Akzeptanz und Umsetzung entscheidend sind. Um Leitlinienennutzern eine Orientierung über das Ausmaß der Berücksichtigung dieser Aspekte zu ermöglichen, werden 4 Klassen von Leitlinien unterschieden (s. Tab. 35.2).

Tab. 35.2: Stufenklassifikation von Leitlinien der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) (nach [8, 23])

Bezeichnung	Charakteristika	Wissenschaftliche Legitimation der Methode	Legitimation für die Umsetzung
S1 Handlungsempfehlungen von Experten	1. Selektierte Entwicklergruppe 2. Keine systematische Evidenzbasierung 3. Keine strukturierte Konsensfindung	gering	gering
S2k Konsensbasierte Leitlinien	1. Repräsentative Entwicklergruppe 2. Keine systematische Evidenzbasierung 3. Strukturierte Konsensfindung	gering	hoch
S2e Evidenzbasierte Leitlinien	1. Selektierte Entwicklergruppe 2. Systematische Evidenzbasierung 3. Keine strukturierte Konsensfindung	hoch	gering
S3 Evidenz- und Konsensbasierte Leitlinien	1. Repräsentative Entwicklergruppe 2. Systematische Evidenzbasierung 3. Strukturierte Konsensfindung	hoch	hoch

35.2 Wie werden evidenz- und konsensbasierte Leitlinien entwickelt?

Für die Autoren von Leitlinien stellt sich die Frage, wie die Aspekte der Evidenz- und Konsensbasierung bei der Leitlinienentwicklung erfüllt und wie die in Kap. 15 ausführlich dargestellten Kriterien von DELBI berücksichtigt werden können. Aus praktischer Perspektive lassen sich 5 Phasen im Leben einer Leitlinie unterscheiden, die schrittweise abgearbeitet bzw. angesprochen werden können: Planung und Organisation, Leitlinienentwicklung, Redaktion und Verbreitung, Implementierung, Evaluierung und Fortschreibung (vgl. Abb 35.1).

35.2.1 Planung und Organisation der Leitlinienentwicklung

Die Erstellung von Leitlinien ist mit erheblichem zeitlichem, personellem und finanziellem Aufwand verbunden. Deshalb bedarf es zunächst einmal einer Auswahl geeigneter Themen, die sich danach richten sollte, dass

die zu erarbeitenden Leitlinien wichtige Fragen der Entscheidungsfindung adressieren und mit ihrer Hilfe ein bestehendes Potenzial zur Verbesserung der Gesundheitsversorgung ausgeschöpft werden kann. Orientierungspunkte für die Priorisierung von Leitlinienthemen sind [7; 13]:

- ▲ Optimierungspotenzial der Versorgungsqualität,
- ▲ Häufigkeit (Prävalenz/Inzidenz) des Versorgungsproblems,
- ▲ Relevante, vermeidbare Unterschiede in der Versorgungspraxis,
- ▲ Krankheitslast,
- ▲ Ethische und soziale Aspekte,
- ▲ Kostenaufwand,
- ▲ Informationsbedarf, z.B. bei neuen Technologien.

Ist ein Thema gefunden, sollte bereits in der Planungsphase der Zielhorizont der Leitlinie skizziert werden. Auf dieser Basis wird das Leitliniengremium einberufen, wobei alle im Anwendungsbereich der Leitlinie an der Versorgung beteiligten Fachgruppen und auch die Patienten, für die die Leitlinie gelten soll, angesprochen werden sollten. Im Programm

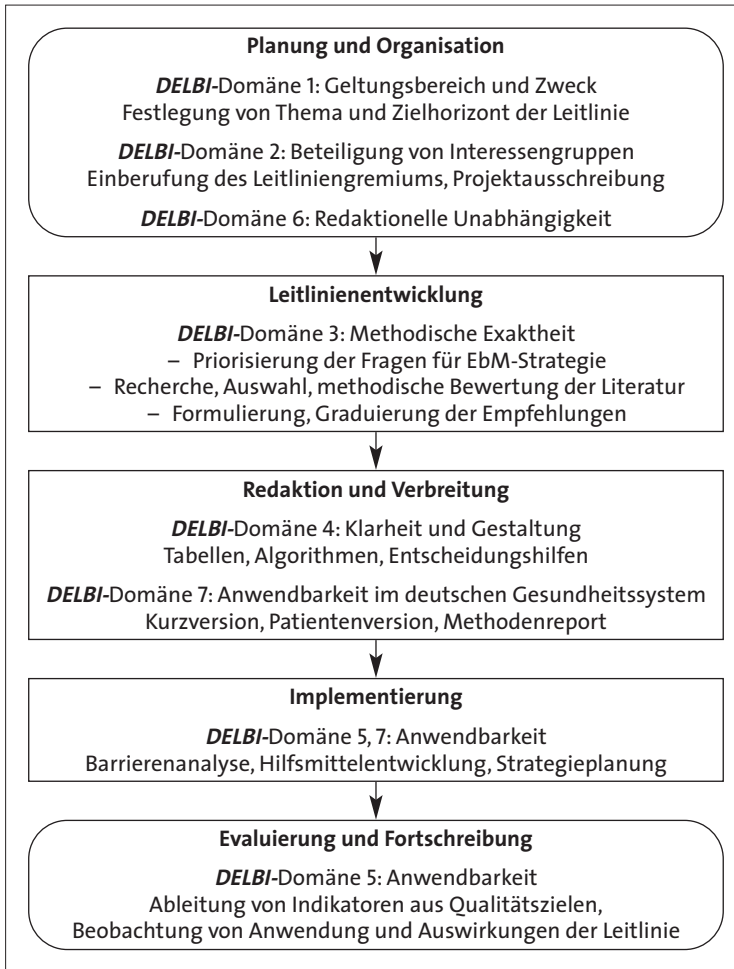


Abb. 35.1: 5 Phasen im Leben einer Leitlinie

für Nationale Versorgungsleitlinien von Bundesärztekammer (BÄK), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) und Kassenärztlicher Bundesvereinigung (KBV) am Ärztlichen Zentrum für Qualität in der Medizin (ÄZQ) [24] und im Leitlinienprogramm der AWMF [23] werden laufende Projekte auch öffentlich ausgeschrieben, um sicherzustellen, dass allen interessierten Gruppen eine Beteiligung und/oder Kommentierung möglich ist (siehe im Internet: www.versorgungsleitlinien.de, www.awmf-leitlinien.de).

Ist das Leitliniengremium bestellt, ist als nächstes die Frage der Finanzierung des Projekts unter Wahrung redaktioneller Unab-

hängigkeit zu klären. Eventuelle Interessenkonflikte von Leitlinien-Autoren sind ebenso darzulegen wie die Finanzierung der Leitlinien-Erstellung.

35.2.2 Leitlinienentwicklung

35.2.2.1 Priorisierung der Fragen für die EbM-Strategie

Im Kontext der Leitlinienarbeit bedeutet Evidenzbasierung konkret die systematische Recherche, Auswahl und Bewertung der Literatur zu relevanten klinischen Fragestellungen. Leitlinien sollen nicht nur einzelne Fragestellungen beantworten, sondern praxisnahe

Empfehlungen und Hilfestellungen für Entscheidungen über ganze Versorgungsabläufe bereitstellen. Klinische Fragestellungen werden von der Leitliniengruppe generiert. Dabei kann sich aus einer einzigen, aus klinischer Perspektive wichtigen Fragestellung eine Fülle suchtauglicher Fragen ergeben. Die Übersetzung der klinischen in suchtaugliche Fragen und die Eingrenzung der EbM-Strategie auf das Notwendige und Machbare erfordern enge Zusammenarbeit und Absprache zwischen Klinikern und Methodikern.

In einem ersten Schritt stellt die Leitliniengruppe fest, welche Fragestellungen durch Empfehlungen allein auf der Basis strukturierter Konsensfindung beantwortet werden sollen. Um Doppelarbeit zu vermeiden, ist im zweiten Schritt eine Überprüfung sinnvoll, welche Fragestellungen durch Adaptation von Empfehlungen bereits existierender Leitlinien beantwortet werden können [7; 24; 42]. International steht mittlerweile eine große Zahl methodisch hochwertiger Leitlinien zur Verfügung [1]. Die Auswahl von Quell-Leitlinien richtet sich in der Praxis nach thematischer Relevanz, Aktualität, Übertragbarkeit auf das deutsche Gesundheitssystem und methodischer Qualität gemäß der Bewertung nach DELBI.

Eine gute Orientierungshilfe für die Priorisierung stellt die Skizzierung des realen Versorgungsablaufs dar. Die Darstellung als Flussdiagramm (Algorithmus) erleichtert die Klärung, über welche Entscheidungs- und Handlungssituationen Unsicherheit besteht, die eine Überprüfung durch systematische Literaturlaufarbeitung nach den Kriterien der EbM erforderlich macht. Ein Praxisbeispiel aus der Nationalen VersorgungsLeitlinie Koronare Herzerkrankung illustriert dieses Vorgehen, ausgehend von der Fragestellung [25]: Wie sollte die Abfolge nicht-invasiver diagnostischer Schritte erfolgen?

35.2.2.2 Recherche, Auswahl und Bewertung der Literatur nach den Kriterien der EbM

Die Systematik der Literaturrecherche und -bewertung ist in Kapitel 20 detailliert beschrieben, deshalb soll hier nur auf die für die Praxis der Leitlinienentwicklung relevanten Aspekte eingegangen werden. Zunächst ist zu überlegen, ob eine Eingrenzung der Recherche auf Studien mit dem hinsichtlich der Fragestellung am besten geeigneten Design möglich ist (z.B. systematische Übersichtsarbeiten von RCT bei therapeutischen Interventionen). In der Realität wird allerdings das Spektrum relevanter Zielgrößen – einschließlich Lebensqualität oder Risiken und unerwünschte Ereignisse – in hochwertigen Studien oft nicht in einem für die Formulierung von Leitlinienempfehlungen ausreichenden Maße erfasst, sodass man auch auf weniger gut angelegte Studien zurückgreifen muss, entsprechend dem Prinzip der systematischen Suche nach der **besten verfügbaren** Evidenz zur Fragestellung [26].

Die Charakteristika, methodische Stärken und Schwächen sowie die Ergebnisse der Studien sollten in Form von Evidenztabellen zusammenfassend dargestellt werden [27]. Dabei wird die Stärke der Evidenz (auch: Level of Evidence, Evidenzgrad) gekennzeichnet, die mit der methodischen Validität der Studien korrespondiert. Die Darlegung der Evidenzstärke soll einen Eindruck von der Verlässlichkeit der für die Formulierung einer Leitlinienempfehlung zugrunde gelegten Studien bzw. deren Ergebnissen vermitteln.

35.2.2.3 Formulierung und Graduierung von Empfehlungen

Nach Analyse und Darlegung der Evidenz werden im nächsten Schritt die konkreten Leitlinien-Empfehlungen formuliert und mit einem Empfehlungsgrad verknüpft. Die Stärke einer Empfehlung korrespondiert mit dem Ergebnis der klinischen Beurteilung der Aussagefähigkeit und Anwendbarkeit der

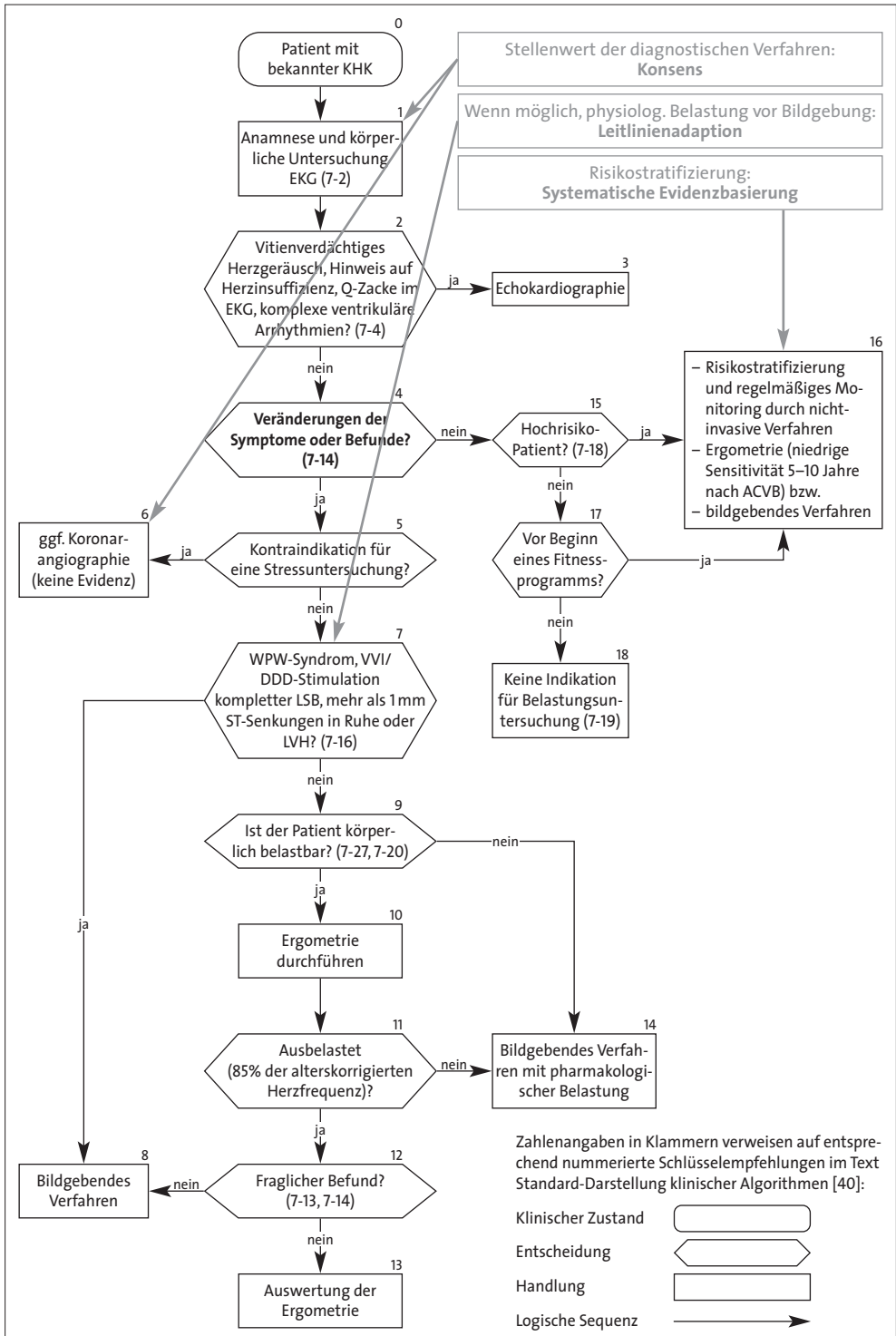


Abb. 35.2: Algorithmus, Beispiel aus Nationaler Versorgungsleitlinie KHK 2006 [25]: Nichtinvasive Diagnostik bei bekannter KHK – kardiologische Versorgungsebene

methodisch aufgearbeiteten Evidenzen, im Englischen auch „considered judgement“ genannt. Diese Beurteilung enthält explizit und implizit **wertende** Elemente und kann als Entsprechung zur Berücksichtigung der internen Evidenz, also der klinischen Erfahrung und des klinischen Urteils, bei Anwendung des Prinzips der Evidenzbasierten Medizin verstanden werden. Leitlinien stellen damit den Brückenschlag zwischen der Welt der externen Evidenz auf der einen Seite und der konkreten klinischen Entscheidungssituation auf der anderen Seite dar.

Konkret sind dabei vor allem folgende Aspekte zu berücksichtigen:

- ▲ Konsistenz der Studienergebnisse,
- ▲ Klinische Relevanz der Studienendpunkte und der Effektstärken (insbesondere im Hinblick auf die Versorgungsziele),

- ▲ Nutzen-Risiko-Abwägung (Ausmaß erwünschter/unerwünschter Effekte, unterschiedliche Evidenzstärken für verschiedene Zielgrößen),
- ▲ Übertragbarkeit der Studien (Möglichkeit der Extrapolation auf erweiterte Patientenzielgruppen, zu erwartende – in der Regel reduzierte – Effektstärken im Alltag),
- ▲ Umsetzbarkeit der Empfehlungen (Bedarf an Strukturen und Ressourcen, Aufwand),
- ▲ Präferenzen der Patienten,
- ▲ Ethische Verpflichtungen [7].

Überwiegend sollte und wird es so sein, dass starke Empfehlungen auf der Grundlage von Evidenz hoher methodischer Qualität ausgesprochen werden. Die genannten Aspekte

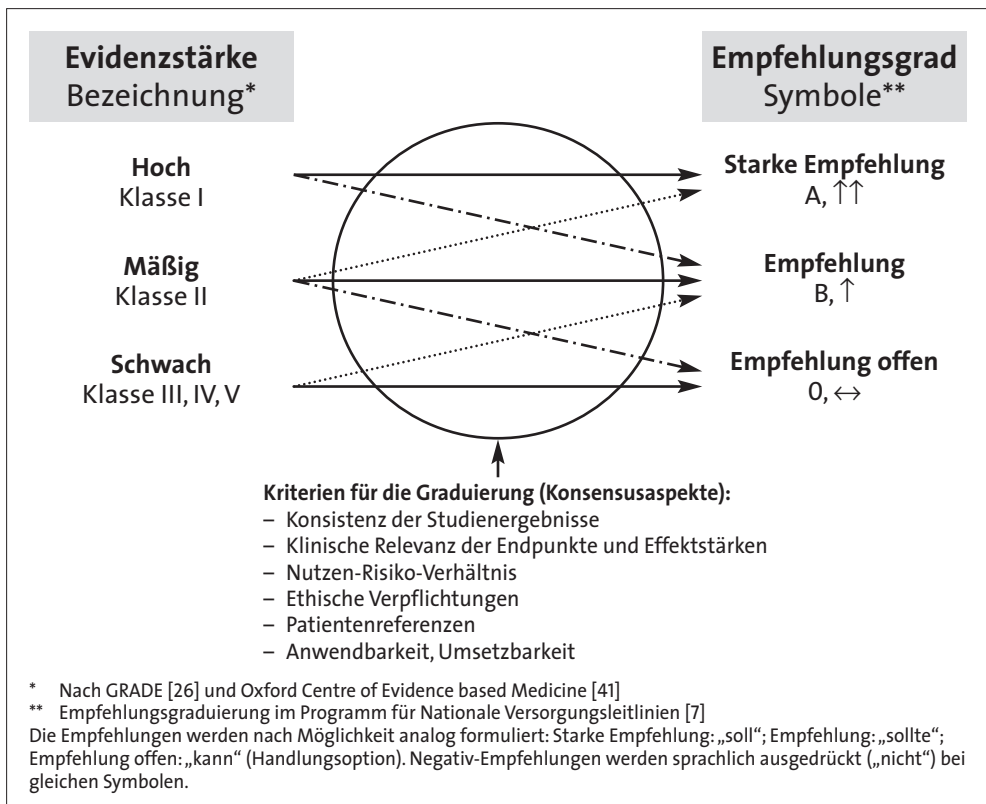


Abb. 35.3: Von der Evidenz zur Empfehlung: Visualisierung der klinischen Beurteilung als Prozess der Kriterien gestützten Konsensusentscheidung [5; 7]

können aber Unterschiede zwischen Evidenzlage und Empfehlung bzw. Empfehlungsstärke begründen (s. Abb. 35.3). Auch wird man in vielen Bereichen der Medizin, in denen Kliniker und Patienten Entscheidungshilfen benötigen, mit mangelnden Evidenzen umgehen müssen [28] – als Beispiel sind hier weite Bereiche der Kinderheilkunde anzuführen.

Die Kriterien für die klinische Bewertung und damit für mögliche Modifikationen des Empfehlungsgrades im Vergleich zur Evidenzstärke sind a priori festzulegen und sollten gezielt in der Leitliniengruppe diskutiert werden. Die Vorgehensweise bei der Formulierung und Graduierung der Empfehlungen sowie die Gründe für Abweichungen von der Evidenzlage müssen nachvollziehbar dargelegt werden, etwa in Form eines der Leitlinie begleitenden Methodenreports [25].

35.2.2.4 Leitlinienentwicklung: Rolle der Konsensfindung

Zu bestimmten Anteilen fließen Erfahrungen, subjektive Einschätzungen und Wertvorstellungen der beteiligten Berufsgruppen und Patienten in alle Kernschritte der Leitlinienentwicklung ein, in die Priorisierung klinischer Fragestellungen, in die Studienauswahl (insbesondere, wenn hochwertige Studien nicht zur Verfügung stehen) und die Gesamteinschätzung der methodischen Qualität der Evidenz, vor allem aber in die Formulierung und Graduierung von Empfehlungen. Sie müssen zu einer nachvollziehbaren Gruppenentscheidung kanalisiert werden, um die Leitlinie vertrauenswürdig zu machen und die Irrtumswahrscheinlichkeit zu minimieren. Der Methodik der Konsensfindung kommt daher ebenso große Bedeutung zu wie der Methodik der Evidenzsynthese. Grundsätzlich sollen deshalb bei der Entwicklung von Leitlinien formale, strukturierte Konsensustechniken mit Beteiligung der potenziellen Anwender und Betroffenen eingesetzt werden [29]. Dazu gehö-

ren Nominaler Gruppenprozess, Delphi-Technik und Konsensuskonferenz (s. Tab. 35.3).

In der Regel werden zwei oder alle drei Techniken bei der Entwicklung einer Leitlinie entsprechend der jeweils anstehenden Aufgabe und der spezifischen Vor- und Nachteile der jeweiligen Verfahren eingesetzt. Sie sind bei richtiger Vorbereitung und Anwendung geeignet, reproduzierbare Ergebnisse zu erreichen sowie unerwünschte Einflüsse und Verzerrungen zu vermeiden, etwa durch Status- und Persönlichkeitsmerkmale der Teilnehmer, persönliche, politische, wirtschaftliche Partikularinteressen und gruppenpsychologische Phänomene [9; 29; 30]. Abstimmungsverläufe und Ergebnisse sind dabei ebenso zu dokumentieren wie Bereiche, in denen kein Konsens erzielt wurde, des Weiteren ggf. auch Minderheitsmeinungen mit entsprechenden Begründungen [25].

35.3 Redaktion und Verbreitung

Die Darstellung der Empfehlungen in der Leitlinie sollte den Ablauf der medizinischen Entscheidungsprozesse nachvollziehen und somit in wichtigen Entscheidungssituationen die benötigten Informationen schnell und eindeutig erkennbar machen. Hilfreich sind dabei Algorithmen, die z.B. diagnostische Abläufe darstellen [31]. Leitlinien sollten möglichst breit disseminiert werden, unter Nutzung von elektronischen und Druckmedien. Um die Berücksichtigung von Empfehlungen in Klinik und Praxis zu erleichtern, bestehen Leitlinien in der Regel aus mehreren Dokumenten. Dazu gehören

- ▲ Langfassungen mit ausführlichen Hintergrundinformationen, Erläuterungen und Evidenztabellen, die die Entstehung der Empfehlungen nachvollziehbar machen,
- ▲ Kurzversionen für die schnelle Information, geeignet zum Abdruck in Fachzeitschriften,

Tab. 35.3: Formalisierte Konsensusverfahren (nach [29])

Bezeichnung	Ablauf	Charakteristika
Nominaler Gruppenprozess	Die Teilnehmer treffen sich persönlich, die Gruppe wird durch einen unabhängigen, geschulten Moderator unterstützt. In einer Stillarbeitsphase nimmt jeder Teilnehmer zunächst schriftlich Stellung, notiert Änderungswünsche und begründet diese. Alle Beiträge werden im Einzelumlaufverfahren registriert, der Gruppe gezeigt, erst danach diskutiert und schließlich zur Abstimmung gebracht.	<ul style="list-style-type: none"> • Sehr hohe Strukturierung der Interaktion • Hohe Beteiligung der einzelnen Teilnehmer • Direkter Austausch in der Gruppendiskussion erleichtert Klärung von Dissens • Eingeschränkte Teilnehmerzahl (15–20)
Delphi-Technik	Die Beiträge werden schriftlich und anonymisiert mittels strukturierter Fragebögen eingeholt, zusammengefasst (oft statistisch) und der Gruppe zurückgemeldet. Die Befragungsrunden werden bis zum Erreichen einer Gruppenantwort (Konsens oder begründeter Dissens) fortgeführt.	<ul style="list-style-type: none"> • Geeignet für größere Teilnehmerzahl • Anonymisierung schützt Einzelbeiträge und unterstützt die Fokussierung auf Inhalte • Hoher Aufwand
Konsensuskonferenz	Im ersten Teil der Konferenz treffen sich die Teilnehmer in themenspezifischen Kleingruppen und erarbeiten gemeinsame Stellungnahmen. Im zweiten Teil der Konferenz werden von Sprechern der Gruppen dem Gesamtplenum die Ergebnisse der Diskussion vorgetragen und die Stellungnahmen zur Abstimmung gebracht. Kleingruppen und Plenum werden durch unabhängige Moderatoren unterstützt.	<ul style="list-style-type: none"> • Geeignet für 80–100 Teilnehmer • Anonymisierte Abstimmung mit elektronischen Systemen möglich • Plenarsitzung eignet sich für erweiterte externe Prüfung und fördert die Akzeptanz • Geringere Möglichkeit zu Interaktion und Strukturierung der Gruppendiskussion

- ▲ Praxishilfen, z.B elektronische Versionen, die in Krankenhaus-Informationssysteme importiert werden können, Fortbildungsmaterialien (CME-Module) und Kitteltaschenversionen,
- ▲ Patientenversionen[7].

Das letzte und vielleicht wichtigste Qualitätskriterium von Leitlinien betrifft die Existenz eines Berichts über die Entstehung der Leitlinie. Dieser so genannte **Leitlinienreport** ist Grundvoraussetzung für Transparenz und Überprüfung der methodischen Qualität der Leitlinienentwicklung. Damit soll eine höhere Sicherheit für den Anwender erzeugt und eine Einschätzung der Wahrscheinlichkeit ermöglicht werden, dass durch die Befolgung von Leitlinien-Empfehlungen die erwünschten gesundheitlichen

und ökonomischen Resultate tatsächlich erzielt werden können [5; 9; 20; 21].

35.4 Implementierung

Medizinische Leitlinien können nur dann wirksam werden, wenn die in ihnen enthaltenen Empfehlungen auch in den ärztlichen Praxisalltag Eingang finden und umgesetzt werden [34]. Diese Erkenntnis fasst in einem Satz die Herausforderung zusammen, der sich die Autoren, Herausgeber und Nutzer medizinischer Leitlinien national und international momentan stellen müssen [32–36]. Voraussetzungen zur Integration der Leitlinien in ärztliches Handeln sind hohe methodische und fachliche Qualität und leichte Verfügbarkeit. Dies allein reicht aber nicht aus.

Tab. 35.4: Implementierung – Definitionen (nach [3; 5; 32–37])

Unter Implementierung versteht man

- konkrete Aktivitäten und Interventionen, um ein Konzept in davon abzuleitende Resultate zu überführen;
- die Einführung eines Konzepts durch geplantes, schrittweises Vorgehen mit dem Ziel, Verhalten zu verändern;
- den Transfer von Erkenntnissen und Empfehlungen aus Leitlinien in individuelles Handeln von Ärzten und anderen Leistungserbringern sowie von Patienten.

Tab. 35.5: Praxisbeispiel: von der Leitlinienempfehlung zum Qualitätsindikator (nach [38, 39])

Leitlinienempfehlung	Qualitätsziel	Indikator (Referenzbereich)	Ist-Wert 2005
Prädiktiver Faktor für die Vorhersage eines Therapieeffektes einer Hormontherapie ist der Östrogen/Progesteronrezeptor-Status. Beim invasiven Mamma-Karzinom ist in der Primärdiagnostik der Hormonrezeptorstatus zu bestimmen.	Immer immunhistochemische Hormonrezeptor-Analyse (Östrogen und Progesteron) bei invasivem Mamma-Karzinom	Anteil Patientinnen mit Hormonrezeptor-Analyse unter allen Patientinnen mit invasivem Mamma-Karzinom > 95%	95,82%
Evidenzstärke 1a, Empfehlungsgrad A	(Prozessqualität)		

Die Implementierung von Leitlinien ist ein aktiver Prozess, der nicht nur Kenntnis, sondern Verhaltensänderungen herbeiführen soll [5].

Ein Teil dieser Implementierung kann von den Leitlinienautoren vorbereitet werden. Dazu gehört die Diskussion potenzieller organisatorischer Barrieren, die eine Umsetzung der in der Leitlinie enthaltenen Empfehlungen erschweren sowie struktureller, personeller und damit möglicher finanzieller Auswirkungen. Auch sollten Vorschläge zur Leitlinienanpassung auf lokale Verhältnisse erarbeitet und Anknüpfungsmöglichkeiten an existierende Projekte des Qualitätsmanagements und der Versorgungsforschung geprüft werden. Die eigentliche Umsetzung der Leitlinie in den Alltag liegt aber letztlich in der Verantwortung der Adressaten. Hierfür steht eine Reihe effektiver Interventionstechniken zur Verfügung, die lokal zu prüfen sind, um eine angemessene Implementierungsstrategie zusammenzustellen [37; 42].

35.5 Evaluierung und Fortschreibung

Eine Leitlinie erfährt ihre Existenzberechtigung letztlich nur dann, wenn sie umgesetzt wird und durch ihre Befolgung die erwarteten gesundheitlichen und ökonomischen Resultate auch tatsächlich erzielt werden können [3; 5]. Zur Überprüfung dieser klinischen Effektivität oder externen Validität von Leitlinien eignen sich Qualitätsindikatoren (s. Tab. 35.5). Darunter versteht man Messgrößen, deren Ausprägung eine Unterscheidung zwischen guten und verbesserungswürdigen Strukturen, Prozessen und/oder Ergebnissen der Versorgung ermöglicht. Qualitätsindikatoren liefern also einerseits Erkenntnisse über die aktuelle Versorgungsqualität. Andererseits helfen sie dabei, den Erfolg der Leitliniendiseminierung und Implementierung zu überprüfen [13; 38]. Wie bei der Leitlinienentwicklung ist auch bei der Formulierung von Qualitätsindikatoren methodische Sorgfalt wichtig [43]. In Analo-

gie zu DELBI wurden daher zur Unterstützung von Autoren und Anwendern auch wichtige Gütekriterien für Qualitätsindikatoren in Form einer Checkliste zusammengefasst [13].

Der Bedarf für die Fortschreibung und Aktualisierung von Leitlinien ergibt sich nicht nur aus der Fortentwicklung wissenschaftlicher Erkenntnisse, sondern auch aus den Ergebnissen der Evaluierung der Leitlinienanwendung.

35.6 Schlussfolgerung – EbM und Leitlinien

Leitlinien unterscheiden sich von Evidenzberichten durch ihre primäre Zielsetzung, für spezielle Versorgungsprobleme explizit ausformulierte Entscheidungshilfen und Handlungsempfehlungen bereitzustellen. Leitlinien werden damit als wichtige Instrumente der Wissensvermittlung und des Qualitätsmanagements im Gesundheitswesen angesehen.

Idealerweise sollen sich die Leitlinienempfehlungen auf Ergebnisse qualitativ hochwertiger klinischer Studien stützen. Zu bestimmten Anteilen werden aber auch Erfahrungen und Wertvorstellungen einfließen, die transparent gemacht und zu einer multiprofessionellen Gruppenentscheidung kanalisiert werden müssen. Der Methodik der Konsensfindung kommt daher bei der Leitlinienentwicklung ebenso große Bedeutung zu wie der Methodik der Evidenzsynthese.

Erst die Anwendung einer Leitlinie in der Klinik oder der Praxis entscheidet aber über ihren Nutzen für Patientinnen und Patienten, für behandelnde Ärztinnen und Ärzte und für das Gesundheitssystem insgesamt. Sie muss durch Implementierung aktiv gefördert und durch die Messung der Leitlinienkonformität und der Versorgungsqualität anhand geeigneter Indikatoren begleitet werden.

Zur Beurteilung des Nutzens einer Leitlinie wird man ihre methodische Qualität, die Richtigkeit der Einführungsentscheidung, die Verbreitung, die Umsetzung und ihre Auswirkungen auf Strukturen, Prozesse und Ergebnisse der Versorgung gemeinsam bewerten müssen. Schließlich stellen Leitlinien nur eine von vielen Möglichkeiten zur Verbesserung der Versorgungsqualität dar.

Literatur

- [1] Ollenschläger G et al., Improving the quality of health care: using international collaboration to inform guideline programmes by founding the Guidelines International Network (G-I-N). *Qual Saf Health Care.* (2004), 13, 455–460
- [2] Sackett DL, Kunz R (1999) Evidenzbasierte Medizin. EBM-Umsetzung und -Vermittlung (Deutsche Ausgabe). Zuckschwerdt, Bern u.a.
- [3] Field MJ et al. (1990) *Clinical Practice Guidelines: Directions for a New Program.* National Academy Press, Washington DC
- [4] Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Beurteilungskriterien für Leitlinien in der medizinischen Versorgung – Beschlüsse der Vorstände der Bundesärztekammer und Kassenärztlicher Bundesvereinigung, Juni 1997. *Dt Arztebl.* (1997), 94, A–2154–5. <http://www.Leitlinienmanual.de> (14.02.07)
- [5] Europarat, Entwicklung einer Methodik für die Ausarbeitung von Leitlinien für optimale medizinische Praxis. Empfehlung Rec (2001) 13 des Europarates und Erläuterndes Memorandum. Deutschsprachige Ausgabe. *Z Arztl Fortbild Qualitätssich.* (2002), 96 Suppl III, 1–60. <http://www.Leitlinienmanual.de> (14.02.07)
- [6] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Ärztliche Zentralstelle Qualitätssicherung (ÄZQ), Das Leitlinien-Manual von AWMF und ÄZQ. *Z Arztl Fortbild Qualitätssich.* (2001), 95 Suppl 1, 1–84. <http://www.Leitlinienmanual.de> (14.02.07)
- [7] Bundesärztekammer (BÄK), Arbeitsgemeinschaft der Wissenschaftlichen Medi-

- zinischen Fachgesellschaften (AWMF), Kassenärztliche Bundesvereinigung (KBV). Nationales Programm für Versorgungs-Leitlinien. Methoden-Report. 3rd ed. <http://www.versorgungsleitlinien.de> (14.02.07)
- [8] Kopp I, Encke A, Lorenz W, Leitlinien als Instrument der Qualitätssicherung in der Medizin. Das Leitlinienprogramm der Arbeitsgemeinschaft Wissenschaftlicher Medizinischer Fachgesellschaften (AWMF). Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz. (2002), 45, 223–233
- [9] Helou A et al., Methodische Standards der Entwicklung evidenz-basierter Leitlinien in Deutschland. Konsens zwischen Wissenschaft, Selbstverwaltung und Praxis. Z Arztl Fortbild Qualitätssich. (2000), 94, 330–339
- [10] Antes G, Die Evidenz-Basis von klinischen Leitlinien, Health Technology Assessments und Patienteninformation als Grundlage für Entscheidungen in der Medizin. Z Arztl Fortbild Qualitätssich. (2004), 98, 180–184
- [11] Lelgemann M et al., Leitlinien. Was haben Ärzte und Patienten davon. Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz. (2005), 48, 215–220
- [12] Hart D (2005) Klinische Leitlinien und Recht: Einleitung und Kommentar. In: Badura B et al., Gesundheitsrecht und Gesundheitswissenschaften. Band 8: Klinische Leitlinien und Recht, 7-16. Nomos, Baden-Baden
- [13] Altenhofen L et al., Beurteilung klinischer Messgrößen des Qualitätsmanagements – Qualitätskriterien und -indikatoren in der Gesundheitsversorgung. Konsenspapier von BÄK, KBV und AWMF. Z Arztl Fortbild Qualitätssich. (2002), 96, 2–15
- [14] Gemeinsamer Bundesausschuss (G-BA). Richtlinie über grundsätzliche Anforderungen an ein einrichtungsinternes Qualitätsmanagement für die an der vertragsärztlichen Versorgung teilnehmenden Ärzte, Psychotherapeuten und medizinischen Versorgungszentren (Qualitätsmanagementrichtlinie vertragsärztliche Versorgung) in der Fassung vom 18. Oktober 2005. <http://www.g-ba.de/cms/upload/pdf/abs5/beschluesse/2005-10-18-QM-Vertragsarzt-RL.pdf> (22.05.06)
- [15] Grimshaw JM et al., Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess. (2004), 8, iii–72
- [16] Thomas L et al., Guidelines in professions allied to medicine. Cochrane Database Syst Rev. (2000), CD000349
- [17] Grimshaw JM et al., Changing physicians' behavior: what works and thoughts on getting more things to work. J Contin Educ Health Prof. (2002), 22, 237–243
- [18] Grol R et al., Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. BMJ. (1998), 317, 858–861
- [19] The AGREE Collaboration, Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care. (2003), 12, 18–23
- [20] Encke A et al., Das Deutsche Instrument zur methodischen Leitlinien-Bewertung (DELBI). Dt Arztebl. (2005), 102, A–1912–3. <http://www.delbi.de> (14.02.07)
- [21] Grimshaw JM, Russell IT, Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet. (1993), 342, 1317–1322
- [22] OECD. Improvement of the performance of health care systems Labour Market/Social Policy-OCC. Papers No. 57. <http://www.oecd.org> (13.12.06)
- [23] Selbmann H-K, Encke A, Leitlinien: Steter Prozess der Aktualisierung. Dt Arztebl. (2005), 102, A–404–5
- [24] Ollenschläger G et al., Nationale Versorgungsleitlinien von BÄK, AWMF und KBV: Hintergrund, Methodik und Instrumente. Med Klin (Munich). (2006), 101, 840–845
- [25] Bundesärztekammer (BÄK), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Kassenärztliche Bundesvereinigung (KBV), Ärztliches Zentrum für Qualität in der Medizin (ÄZQ). Nationale Versorgungsleitlinie Chronische Koronare Herzkrankheit. <http://www.versorgungsleitlinien.de/themen/khk> (08.07.06)
- [26] Atkins D et al., Grading quality of evidence and strength of recommendations. BMJ. (2004), 328, 1490–1497
- [27] Schneider M et al. (2005) Management der frühen rheumatoiden Arthritis. Interdisziplinäre Leitlinie. Steinkopff Verlag, Darmstadt
- [28] Neises G, Windeler J, Wie viel ist „evidenzbasiert“? Eine Übersicht zum aktuel-

- len Forschungsstand. Z Arztl Fortbild Qualitatssich. (2001), 95, 95–104
- [29] Black N et al., Consensus development methods: a review of best practice in creating clinical guidelines. J Health Serv Res Policy. (1999), 4, 236–248
- [30] Koller M, Beiträge der Sozialpsychologie zur Analyse und Lösung von Problemen im deutschen Gesundheitssystem – das Beispiel Leitlinien. Z Sozialpsych. (2005), 36, 47–60
- [31] Schoenbaum SC, Gottlieb LK, Algorithm based improvement of clinical quality. BMJ. (1990), 301, 1374–1376
- [32] Kopp I et al., Zur Empirie hochwertiger Leitlinien im System der Arbeitsgemeinschaft Wissenschaftlicher Medizinischer Fachgesellschaften (AWMF): gibt es sie und wie viele? Dtsch Ges Chir-Mitteil. (2005), 21–29
- [33] Kopp I (2005) Wege zur Qualitätsentwicklung: Implementierung von Leitlinien. In: Badura B, Iseringhausen O, Wege aus der Krise der Versorgungsorganisation – Beiträge aus der Versorgungsforschung, 255–266. Huber, Bern
- [34] Ollenschläger G et al. (2005) Qualität und Akzeptanz medizinischer Leitlinien in Deutschland – Bestandsaufnahme Mai 2004. In: Hart D, Klinische Leitlinien und Recht, 17–39. Nomos-Verl.-Ges., Baden-Baden
- [35] Grol R, Personal paper. Beliefs and evidence in changing clinical practice. BMJ. (1997), 315, 418–421
- [36] Grimshaw J, Eccles M, Tetroe J, Implementing clinical guidelines: current evidence and future implications. J Contin Educ Health Prof. (2004), 24 Suppl 1, S31–S37
- [37] Selbmann H-K, Kopp I, Implementierung von Leitlinien in den Versorgungsalltag. Psychiatrie. (2005), 2, 33–38
- [38] Bundesgeschäftsstelle Qualitätssicherung (BQS) (2006) Qualität sichtbar machen. BQS-Qualitätsreport 2005; im Auftrag des Gemeinsamen Bundesausschusses nach § 91 SGB V. Bundesgeschäftsstelle Qualitätssicherung, Düsseldorf
- [39] Deutsche Krebsgesellschaft e.V., Informationszentrum für Standards in der Onkologie (2004) Interdisziplinäre S3-Leitlinie für die Diagnostik, Therapie und Nachsorge des Mammakarzinoms der Frau. Zuckschwerdt, München
- [40] Society for Medical Decision Making Committee on Standardization of Clinical Algorithms, Proposal for clinical algorithm standards. Med Decis Making (1992), 12 (2), 149–154
- [41] Philipps B et al., Levels of evidence and grades of recommendations. Oxford: Oxford Centre of Evidence-Based Medicine. (2001). http://www.cebm.net/levels_of_evidence.asp (08.07.05)
- [42] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), PMV forschungsgruppe, Leitliniengruppe Hessen, Handbuch zur Entwicklung regionaler Leitlinien. Berlin: ÄZQ (2006). <http://www.Leitlinienmanual.de> (14.02.07)
- [43] Geraedts M et al., Beurteilung der methodischen Qualität klinischer Messgrößen. Z Arztl Fortbild Qualitatssich. (2002), 96, 91–96

36 Systemsteuerung im Rahmen des SGB V: der Gemeinsame Bundesausschuss

Bernhard Gibis, Regina Klakow-Franck, Nicole Schlottmann, Johannes Bruns

Es entspricht den Grundprinzipien des deutschen Sozialversicherungssystems, dass der Rahmen der Gesundheitsversorgung durch den Gesetzgeber festgesetzt und die Ausgestaltung der Details bzw. Umsetzung der Maßnahmen den so genannten Partnern der Selbstverwaltung überlassen wird. In diesem Rahmen hat sich seit Verkündung der „Kaiserlichen Botschaft“ (1881) bzw. der Verabschiedung des Gesetzes „betreffend die Krankenversicherung der Arbeiter“ (1885) in der Selbstverwaltung der Krankenversicherung ein zentrales Lenkungs- und Steuerungsgremium herausgebildet, das insbesondere 2004 in Form des Gemeinsamen Bundesausschusses (G-BA) mit umfassenden Befugnissen hinsichtlich Zuständigkeit und Regelungskompetenz ausgestattet wurde.

Zunächst wurde 1913 – nach der grundsätzlichen Entscheidung, kein steuerfinanziertes, also staatliches Gesundheitssystem zu etablieren sowie nach erbitterten Auseinandersetzungen zwischen Ärzten und Krankenkassen – der Zentrallausschuss der Ärzte und Krankenkassen etabliert. Dieser wurde 1923 nach erneuten Streikwellen in einen Reichsausschuss mit weitergehenden Zuständigkeiten überführt. Walter Spielhagen, ein zuständiger Beamter des damaligen Reichsarbeitsministeriums, fasste seinerzeit zusammen: „In der aus Erfahrung gewonnenen Erkenntnis, dass auf dem Gebiete der Beziehungen zwischen Ärzten und Krankenkassen der Reichstag nicht die geeignete Stelle zur Lösung der zahlreichen in Betracht kommenden Einzelfragen ist, hat der Gesetzgeber auf diesem kleinen, aber wichtigen Gebiete einen Teil seiner eigenen Funktionen auf ein paritätisch

besetztes Organ übertragen. Indem er diesem Organ bestimmte öffentlich-rechtliche Funktionen zuweist, macht er die beteiligten Verbände aus bloßen Objekten zu Subjekten der Gesetzgebung.“

Dieses Grundprinzip hat sich offensichtlich bewährt und wechselvolle Zeitläufte überstanden. Es war allerdings nicht kompatibel mit der Auffassung eines Staates, der die detaillierte Regelung der Gesundheitsversorgung für sich selbst reklamiert, wie dies 1933–1945 in ganz Deutschland und anschließend auf dem Gebiet der DDR der Fall war.

Der nach den Wirren der Nazi-Zeit in der Bundesrepublik eingerichtete Bundesausschuss der Ärzte und Krankenkassen wurde mit nahezu jeder neuen Gesetzesreform ausgebaut und zu einem zentralen Steuerungsgremium der Gesetzlichen Krankenversicherung (GKV) ausgestaltet. War zunächst die Zuständigkeit ausschließlich auf die ambulante Versorgung beschränkt, so wurde der stationäre Sektor im Jahre 2000 durch die Einrichtung des Koordinierungsausschusses und des Ausschusses Krankenhaus mit einbezogen. Mit der Etablierung des Gemeinsamen Bundesausschusses (2004) wurde diese sektorenübergreifende Ausrichtung weiter betont. Träger des G-BA sind nach Vorgabe des Gesetzgebers die Spitzenverbände der Gesetzlichen Krankenkassen, die Deutsche Krankenhausgesellschaft und die Kassenärztliche Bundesvereinigung. Nicht zuletzt aufgrund zunehmender Zweifel an der Legitimität eines Gremiums, das weitreichende Entscheidungen zur Gesundheitsversorgung von über 70 Millionen Bürgern trifft, die in

anderen Ländern durch das Parlament vorgenommen werden [1], wurde der Kreis der Akteure im G-BA um Patientenvertreter erweitert.

Die Arbeit des G-BA konzentriert sich auf Leistungs- und Zulassungsfragen im Rahmen der GKV und – durch die letzten Gesetzgebungen verstärkt – auch auf den Aspekt der Versorgungsqualität. Der G-BA erlässt Richtlinien, die als untergesetzliche Normen für alle im Rahmen der GKV am Gesundheitswesen Beteiligten unmittelbar verbindlich sind. Seine gesetzliche Grundlage findet er im Sozialgesetzbuch 5 (SGB V), das detailliert seine Aufgaben und Funktionen festschreibt. Dem Bundesministerium für Gesundheit (BMG) kommt dabei eine Aufsichtsfunktion zu, die sich traditionell primär auf den formalen Ablauf der Beschlussfassung und weniger auf deren Inhalte konzentriert. Dieser Weg wird allerdings, wie Beanstandungen des BMG zu mehreren Entscheidungen des G-BA zeigen (z.B. zur Sondernahrung und zur Protonentherapie), immer mehr verlassen.

36.1 Entscheidungsfindung im Gemeinsamen Bundesausschuss

Grundlage der Entscheidungsfindung im Gemeinsamen Bundesausschuss ist der vom Gesetzgeber vorgegebene Rahmen, der mit unbestimmten Rechtsbegriffen die Charakteristika derjenigen Leistungen beschreibt, die in der Gesetzlichen Krankenversicherung vorgehalten werden müssen. Das begriffliche Spektrum reicht dabei von „ausreichend“, „zweckmäßig“, „wirtschaftlich“ bis hin zu „medizinisch notwendig“, „nützlich“ und „erforderlich“. Leistungen der GKV müssen zudem dem „allgemein anerkannten Stand der wissenschaftlichen Erkenntnis entsprechen“. Eine weitere Hinterlegung dieser Begriffe wurde nicht vorgenommen. Es bleibt letztlich den Akteuren einer jeweiligen Generation überlassen, diese Begriffe mit Leben zu füllen.

Die Interpretation dieser Begriffe ist einem beständigen Wandel unterworfen: Wäh-

Plenum	Ärztliche Angelegenheiten	Vertragsärztliche Versorgung	Vertragszahnärztliche Versorgung	Krankenhausbehandlung
Geschäftsordnung und Finanzen	Disease-Management-Programme	Ärztliche Behandlung	Bedarfsplanung	Methodenbewertung
Verfahrensordnung	Ambulante Behandlung im Krankenhaus	Arbeitsunfähigkeit	Festzuschüsse/ Zahnersatz	Externe stationäre Qualitätssicherung
Patienteninformation	Empfehlungen zur Qualitätssicherung (in Gründung)	Arzneimittel	Kieferorthopädie	Sonstige stationäre Qualitätssicherung
		Bedarfsplanung	Neue Untersuchungs- Behandl.methoden	
		Familienplanung	Qualitätsbeurteilung/ -sicherung	
		Hausliche Krankenpflege	Richtlinien	
		Heil- und Hilfsmittel		
		Krankenhaus- behandlung/ Krankentransport		
		Prävention		
		Psychotherapie (§ 91 Abs. 5 Satz 2)		
		Qualitätsbeurteilung/ -sicherung		
		Rehabilitation		
		Soziotherapie		

Abb. 36.1: Struktur und Unterausschüsse des Gemeinsamen Bundesausschusses, Stand Oktober 2006

rend beispielsweise früher als physiologische Alterungsprozesse erachtete Gesundheitszustände nicht als Gegenstand von Leistungen der GKV gesehen wurden (z.B. die weibliche Menopause), so hat sich hier ein grundlegender Wandel vollzogen. Gleiches gilt auch für die Begrifflichkeit der Lebensqualität.

Entscheidungen im G-BA werden in einem Spannungsfeld getroffen, das von den häufig widerstrebenden Interessen der beteiligten Akteure dominiert wird. Entscheidungen können dabei durch verschiedene Grundhaltungen determiniert werden, so z.B. durch Instinkt („gesunder Menschenverstand“), Gewohnheit („Paradigma“) oder verbandliche Interessen („finanzielle Anreize“).

In Anbetracht des Drucks rückläufiger Ressourcen und wettbewerblicher Aufstellung ganzer Versorgungssektoren ist es deshalb von entscheidender Bedeutung, den Prozess der Entscheidungsfindung so zu gestalten, dass er die Akzeptanz aller Beteiligten findet. Nicht immer können Entscheidungen so klar und sinnfällig sein, dass alle Beteiligten das Ergebnis der Entscheidungsfindung anerkennen. Auch wenn eine Einzelfrage zuungunsten eines der Akteure entschieden wird, so wird dieses Ergebnis nur dann akzeptiert, wenn der dort hinführende Entscheidungsprozess als fair und legitim erlebt wird (sog. prozedurales Einverständnis nach [2]).

Grundsätze eines fairen Entscheidungsprozesses hat Leventhal [3] zusammengefasst:

- ▲ konsistente Prozeduren,
- ▲ die unvoreingenommen durchgeführt werden,
- ▲ auf der Basis präziser Informationen,
- ▲ mit der Möglichkeit der Korrektur,
- ▲ unter Berücksichtigung der Interessen der beteiligten Parteien,
- ▲ gemäß ethischen und moralischen Standards.

Informationen zu medizinischen Fragestellungen zur Politikberatung wurden traditionell durch Expertengremien in den Vorläuferorganisationen des G-BA immer dann bereitgestellt, wenn Wissen zur Lösung fachspezifischer Fragen erforderlich war. Seit der Etablierung der Gesetzlichen Krankenversicherung war der Expertenrat traditionell Grundlage der „präzisen Informationen“ und sowohl individuell als auch als Organisation (z.B. als wissenschaftliche Fachgesellschaft) als Wissensquelle in die Politikberatung integriert. Dies kann mit mehreren Vorteilen verbunden sein: Experten können schnell eine Einschätzung über ihr Fachgebiet abgeben, sind in der Regel über Jahre mit Entwicklungen vertraut und können Hintergrundinformationen, die nicht publiziert werden, weitergeben. Zudem repräsentieren Experten oft Interessengruppen, die von späteren Entscheidungen betroffen sein können. Ihre Integration in Entscheidungsabläufe kann deshalb auch im Hinblick auf die spätere Umsetzung von Vorteil sein [4].

Der Nimbus der mündlichen Expertenanhörung hat in den letzten Jahren jedoch wesentlich an Bedeutung verloren: Experten verfügen nicht mehr allein über relevantes Wissen zu den Charakteristika einer Methode, selbst die gruppenspezifische Repräsentativität der vorgetragenen Fakten und Meinungen ist nur schwerlich überprüfbar. Da Experten die zu prüfenden Verfahren teilweise selbst entwickelt haben oder anwenden, kann die Sicht darauf entweder durch persönliche Interessen oder durch enge Verbindungen mit der die Forschung finanzierenden Industrie überlagert werden. Beklagt wird zudem, dass manche Argumente nicht faktenbasiert vorgetragen werden, sondern sich auf nicht minder wirkungsvolle Hilfsargumente beziehen (z.B. „Wenn ich krank wäre, würde ich mich nur damit behandeln lassen!“).

36.1.1 Evidence Based Health Care (EBHC)

Im angelsächsischen Umfeld hat sich deshalb, in Analogie zur Evidence based medicine, die Vorgehensweise der „Evidence based health care“ (EBHC) etabliert [5]. Während bei der EbM die persönliche Erfahrung des Arztes, die Wünsche des Patienten und die Erkenntnisse aus kontrollierter wissenschaftlicher Erfahrung (Evidenz) eine immer wiederkehrende Trias bilden, werden Entscheidungen nach den Prinzipien der evidenzbasierten Gesundheitsversorgung vor dem Hintergrund der vorhandenen Evidenz, der zur Verfügung stehenden Ressourcen und der Wertevorstellungen, wie sie in jedem Gesundheitssystem implizit und explizit vorhanden sind, gebildet (s. Abb. 36.2). Experten haben in diesem Gebilde nach wie vor eine wichtige, wenn auch deutlich relativierte Funktion. Von besonderer Bedeutung sind dabei aggregierte, zusammenfassende Produkte der Evidenzbasierten Medizin wie Leitlinien oder systematische Übersichtsarbeiten. Speziell für die Erarbeitung entscheidungsrelevanter Unterlagen haben sich HTA-Berichte etabliert, die das verfügbare Wissen, eine Technologie oder Maßnahme betreffend, zusammenfassen und so Entscheidungsträger in die Position einer „informierten“ Entscheidung versetzen sollen (s. Kap. 2 und 16).

In den staatlich dominierten Gesundheitswesen wurden in den letzten Jahren kontinuierlich Elemente der evidenzbasierten Gesundheitsversorgung in den Entscheidungsablauf integriert, insbesondere in Großbritannien. Unter den Sozialversicherungssystemen haben hierbei die Niederlande eine führende Rolle eingenommen [6].

36.1.2 EBHC in den Vorläuferorganisationen des G-BA

In den Vorläuferorganisationen des G-BA wurden schon frühzeitig Elemente der Evidenzbasierten Medizin in die Entscheidungsfindung integriert. Schon im Jahre 1990 wurden die Richtlinien des Arbeitsausschusses „Ärztliche Behandlung“ für die Bewertung von vorliegenden wissenschaftlichen Untersuchungen mit „Evidenzstufen“ versehen (Beschluss des Bundesausschusses vom 4.12.1990).

Mit der Veröffentlichung der Verfahrensrichtlinie vom 1.1.1998 im Bundesanzeiger wurden weitere Elemente der Evidence based health care etabliert: Entscheidungen des Unterausschusses wurden Berichte zugrunde gelegt, die eine Aufbereitung der wissenschaftlichen Evidenz, die Ergebnisse einer schriftlichen Sachverständigenanhörung und den

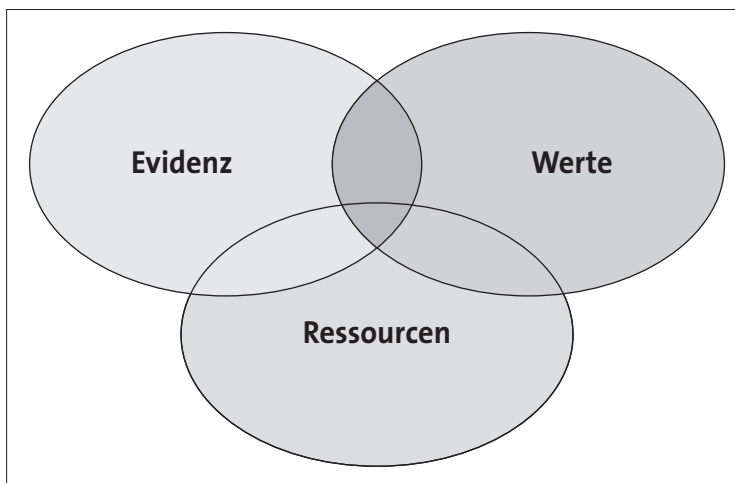


Abb. 36.2: Entscheidungsgrundlagen der Evidence Based Health Care [5]

Tab. 36.1: Vor- und Nachteile der evidenzgestützten Entscheidungsfindung im Arbeitsausschuss „Ärztliche Behandlung“ (nach [9])

Vorteile	Nachteile
Transparente, reproduzierbare Entscheidungsfindung	Zeit- und Ressourcenaufwand
Umfassende Bewertung der relevanten wissenschaftlichen Unterlagen	Rasche Inaktualität
Relative Expertenunabhängigkeit	Fokussierung auf einzelne Technologien und Maßnahmen (und nicht auf Versorgungsabläufe)
Erhöhte Gerichtsfestigkeit	Kenntnisse der evidenzbasierten Bewertung noch ungenügend verbreitet
Gemeinsame Beratungsbasis der Selbstverwaltungspartner	

Status der zu beratenden Technologie im Ausland berücksichtigen [7].

Die Erfahrungen mit diesem Verfahren zeigten, dass es möglich wurde, gerichtsfeste Ergebnisse zur Bewertung einzelner, oft höchst kontrovers diskutierter Behandlungsverfahren zu erzielen. Mit der Veröffentlichung der Berichte, die zudem Daten zum Beratungsverfahren enthalten, wurde eine Transparenz erreicht, wie sie bislang nicht im Bundesausschuss vorhanden war. Befürchtungen, dass diese Transparenz vor Gerichten zum Nachteil gereichen oder aber Ergebnisse von den Medien aus dem Zusammenhang gerissen würden, bestätigten sich nicht. Die Ablehnung von Verfahren hat zudem nicht selten zur bislang fehlenden wissenschaftlichen Forschung geführt: Im Falle der Akupunktur, der Balneophototherapie oder aber der Extrakorporalen Stoßwellentherapie wurden hochwertige wissenschaftliche Untersuchungen aufgelegt, die die zum Entscheidungszeitpunkt vorhandene Wissenslücke schließen sollen [8].

Deutlich wurde aber auch, dass die Evidenzbasierung keinen Algorithmus darstellt, der automatisiert Entscheidungsfindung vorge- oder abnimmt. Im Gegenteil: Bei der Bestimmung der medizinischen Notwendigkeit eines Verfahrens oder auch der Evidenz sind letztlich immer Entscheidungen mit

Augenmaß erforderlich. Bei fehlenden Therapieoptionen zur Behandlung einer schwerwiegenden Erkrankung muss ein anderes Evidenzniveau herangezogen werden als für eine Erkrankung, für die schon eine Vielzahl von effektiven Therapieoptionen zur Verfügung steht.

Immer wiederkehrende Fragen sind die nach der Bewertung der komplementärmedizinischen Verfahren (z.B. Rolle der homöopathischen Erstanamnese), der Diagnostik und Therapie bei Patienten in ausweglosen Situationen (z.B. hyperbare Sauerstofftherapie oder Protonentherapie bei Karzinomkrankungen) und nicht zuletzt die Auswirkungen des evidenzbasierten Verfahrens auf die Einführung („Segel oder Anker“?) von Innovationen (z.B. Verfahren zur Ablation des Endometriums) [9].

36.2 EBHC im G-BA

Die Erfahrungen mit der Anwendung evidenzbasierter Grundsätze in der Entscheidungsfindung sind sowohl in die Gesetzgebung (beispielsweise bei der Beschreibung der Disease-Management-Programme) als auch in die Arbeit des mit dem Gesundheitssystemmodernisierungsgesetz (GMG) in seiner Bedeutung erheblich aufgewerteten GBA

eingeflossen. Wie bereits erwähnt, lässt sich der G-BA grob in drei Bereiche gliedern: einen stationären, einen ambulanten und einen sektorenübergreifenden Anteil.

Während in der ambulanten Versorgung die Bewertung von einzelnen Leistungen (z.B. therapeutischer und diagnostischer ärztlicher Maßnahmen in Therapie, Prävention oder Schwangerenvorsorge, Arzneimittel), Versorgungsabläufen (z.B. Screening-Programmen) und die Richtlinien zur Versorgungsqualität im Vordergrund stehen, sind dies im stationären Sektor insbesondere die Bewertung von Technologien zur Anwendung im Krankenhaus und die Entwicklung von Anforderungen an die Qualitätssicherung.

Hierunter fallen die externe Qualitätssicherung mithilfe des Vergleichs von Indikatoren, die Etablierung von Mindestmengen insbesondere für Operationsleistungen für die Leistungserbringung im Krankenhaus sowie die Festlegung der Anforderung an Qualitätsberichte der Krankenhäuser.

Die sektorenübergreifenden Aufgaben beziehen sich auf die Entwicklung von Empfehlungen für Disease-Management-Programme sowie die Benennung von Indikationen für seltene und schwere Erkrankungen, die der ambulanten Behandlung im Krankenhaus bedürfen (§ 116 b SGB V).

Die benannten Themen zeigen die Reichweite der Richtlinien des G-BA an: so bedarf die Neuaufnahme von Leistungen in den ambulanten Leistungskatalog der GKV der vorherigen Beschlussfassung des G-BA (Erlaubnisvorbehalt), für den stationären Bereich können Richtlinien die Leistungserbringung untersagen (Verbotsvorbehalt). Zu berücksichtigen ist dabei das Leistungsspektrum der GKV, das die Regelversorgung aller Versicherter sicherstellen soll. Methoden, die noch in Forschung begriffen sind oder die nicht den Anspruch der medizinischen Notwendigkeit erfüllen (z.B. sog. Wellness- oder Lifestyle-Methoden) sind nicht Gegenstand der Entscheidungsfindung des G-BA.

36.2.1 Verfahrensordnung des Gemeinsamen Bundesausschusses [10]

Zur Vereinheitlichung der Vorgehensweise hat der Gemeinsame Bundesausschuss eine Verfahrensordnung erstellt, die das Beratungsverfahren standardisiert.

Die Verfahrensordnung bezweckt transparente und rechtssichere Entscheidungen, die dem allgemein anerkannten Stand der medizinischen Erkenntnisse und deren Hinweisen zu Qualität, Versorgungsaspekten von Alter, biologischem und sozialem Geschlecht sowie lebenslagenspezifischen Besonderheiten entsprechen, die berechtigten Interessen der Betroffenen angemessen berücksichtigen und das Gebot der Wirtschaftlichkeit im Sinne des §12 Abs. 1 SGB V beachten. (Verfahrensordnung des G-BA vom 25.9.2005)

Die Verfahrensordnung in der Fassung vom 1.10.2005 (zuletzt in überarbeiteter Version in Kraft getreten am 7.7.2006) umfasst allgemeine Verfahrensbestimmungen wie den Anwendungsbereich, das Beratungsverfahren in den Unterausschüssen mithilfe von Arbeitsgruppen einschließlich der Beschlussfassung und Veröffentlichung nach Vorlage beim BMG. Diese Anforderungen treffen prinzipiell auf alle Unterausschüsse des G-BA zu. Grundsätzlich gilt (§ 5 Abs. 2 der Verfahrensordnung), dass der G-BA „den allgemein anerkannten Stand der medizinischen Erkenntnisse auf der Grundlage der Evidenzbasierten Medizin“ ermittelt. Eine Sonderregelung wurde ebenfalls für die Richtlinienbeschlüsse nach § 116b SGB V getroffen, die hochspezialisierte Leistungen, seltene Erkrankungen sowie Erkrankungen mit besonderen Erkrankungsverläufen betreffen und damit die Grundlage für Verträge zwischen Krankenkassen und zugelassenen Krankenhäusern darstellen. Gegenstand der Verfah-

rensordnung ist ebenfalls die Organisation des gesetzlich vorgeschriebenen Verfahrens zur Ermittlung von Stellungnahmen, die in die Entscheidungsfindung einzubeziehen sind. Da die Entscheidungen des G-BA nicht selten erhebliche (auch finanzielle) Auswirkungen auf den Versorgungsalltag haben, wurde zuletzt eine wie bei der Erstellung von Leitlinien übliche Regelung zur Offenlegung von potenziellen Interessenskonflikten ergänzt.

Sonderregelungen wurden für die Bewertung medizinischer Methoden im Krankenhaus (§ 137c SGB V) und für neue und bereits erbrachte Methoden im ambulanten Bereich (§ 135 Abs. 1 SGB V) etabliert. Hierzu gehören ärztliche Untersuchungs- und Behandlungsmethoden einschließlich der Methoden der Richtlinien-Psychotherapie, der zahnärztlichen Behandlung, der Maßnahmen zur Früherkennung von Krankheiten, der ärztlichen Betreuung bei Mutterschaft und Schwangerschaft sowie der medizinischen Maßnahmen zur Empfängnisverhütung, zum Schwangerschaftsabbruch und zur Sterilisation.

Die Verfahrensordnung regelt die Antragsberechtigung (prinzipiell für die in dem jeweiligen Plenum beteiligten Bänke und Patientenvertreter), die Form der Antragstellung und die Veröffentlichung des Bera-

tungsthemas. Die nun folgende Arbeit wird durch Themenarbeitsgruppen, die bei sektorenübergreifenden Themen mit allen beteiligten Bänken zu besetzen sind, in Angriff genommen. Dabei werden die für die Bewertung der medizinischen Methoden vorhandenen Unterlagen nach Evidenzstufen geordnet und hinsichtlich ihrer Qualität ausgewertet. Insbesondere zu Fragen der Früherkennung werden die gesetzlichen Anforderungen operationalisiert und mit konkreten Fragestellungen hinterlegt.

Das systematische Zusammentragen aller relevanten Informationen im Rahmen der Verfahrensordnung kann nicht die abwägende Entscheidung der Beteiligten ersetzen. Die Verfahrensbeteiligten haben nach kontroverser Diskussion und letztlich nach Maßgabe durch das BMG einen *modus vivendi* gefunden, der jetzt durch gelebtes Handeln in der täglichen Arbeit umgesetzt wird.

Es entspricht den Prinzipien des wissenschaftlichen Erkenntnisgewinns, dass die gewonnenen Ergebnisse je nach Art der Erkenntnis mit einem unterschiedlichen Maß an Unsicherheit einhergehen, ob der beobachtete Effekt tatsächlich der Realität entspricht. Gut belegt ist beispielsweise die Überschätzung von Therapieeffekten in retrospektiven Untersuchungen oder Expertenmeinungen. Als gegenwärtiger Goldstandard wird deshalb

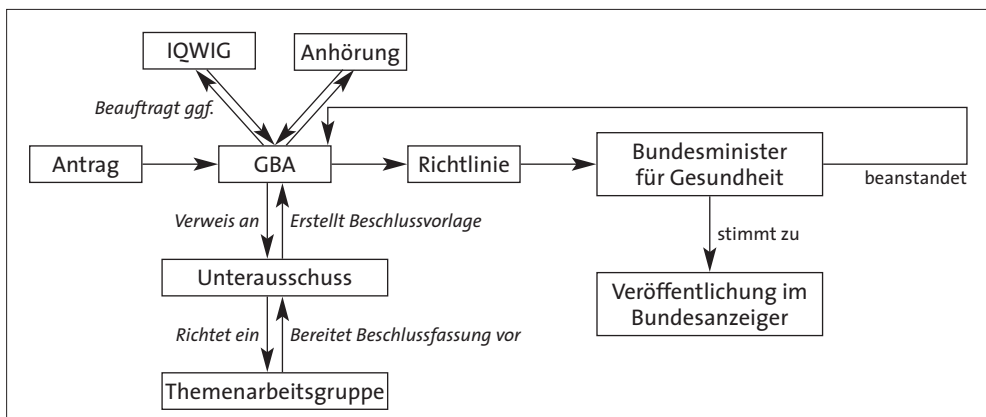


Abb. 36 .3: Prinzipieller Beratungsablauf im G-BA am Beispiel § 91 Abs. 5 SGB V

die vergleichende Untersuchung bei zufälliger Gruppenverteilung der Teilnehmer in eine Standard- und eine Interventionsgruppe angesehen.

Die Verfahrensrichtlinie trägt dieser Tatsache Rechnung und bewertet im Einklang mit den Grundsätzen der Evidenzbasierten Medizin die Ergebnisse solcher, in ihrer umfassendsten Form randomisiert kontrolliert genannter Studien, als höherwertig als die anderer Studienformen. Mit der Hervorhebung dieser Studienart wird auch das Prinzip des kontrollierten Vergleichs als Grundlage des Erkenntnisgewinns hervorgehoben.

Die randomisierte kontrollierte Studie ist jedoch nicht allen Fragestellungen angemessen. Nebenwirkungen und Risiken werden häufig besser in langjährigen Kohortenstudien mit großer Teilnehmerzahl erfasst, bei schwerwiegenden Erkrankungen ohne therapeutische Alternative können auch Erkenntnisse aus Studien niedrigerer Evidenzklasse konkludent sein.

Bei jeder dieser Abweichungen ist jedoch zu berücksichtigen, dass das Maß an Unsicherheit, die tatsächliche Wirksamkeit und auch Gefährdung betreffend, bei Erkenntnissen aus niedrigeren Evidenzstufen potenziell steigt. Es bleibt also einem Entscheidungsgremium überlassen, diese Abwägung unter Berücksichtigung aller vorliegenden Erkenntnisse und Informationen vorzunehmen. Wenn auch die Evidenzbasierte Medizin den Rahmen der Argumentation versachlichen kann, bleibt die Entscheidungsfindung ein schwieriges Unterfangen, wie dies auch die Kontroversen der Akteure im G-BA hinreichend belegen.

Beispiel der Einführung einer Therapie auf unsicherer Evidenzlage ist die Hochdosis-Chemotherapie beim Mammakarzinom. Für diese wurde eine Fallpauschale in die GKV eingeführt, bevor randomisierte kontrollierte Untersuchungen zeigen konnten, dass die Hoch-

dosischemotherapie zwar die Leidenszeit, nicht aber die Überlebenszeit der Patientinnen verlängert. Erste Untersuchungen auf niedrigem Evidenzniveau, die als Begründung für die Einführung herangezogen wurden, waren zudem offensichtlich manipuliert. Nach dem Vorliegen der Ergebnisse randomisierter kontrollierter Studien wurde dieser Ansatz wieder verlassen.

Von wesentlicher Bedeutung ist nicht zuletzt die Bewertung der gefundenen Ergebnisse und insbesondere der medizinischen Notwendigkeit vor dem Hintergrund des Versorgungskontexts der GKV. All dies impliziert, dass die Entscheidungsfindung auf der Grundlage der Verfahrensordnung ein abwägender Prozess ist, der ergebnisoffen durchzuführen und nicht durch Automatismen zu ersetzen ist.

Liste von Entscheidungen, die auf der Grundlage der Verfahrensordnung und von Vorläufern des G-BA getroffen werden.

- ▲ Osteodensitometrie als diagnostische Maßnahme im Rahmen der kurativen Behandlung gemäß § 135 Abs. 1 SGB V
- ▲ Positronenemissionstomographie (PET)
- ▲ Kombinationsverfahren PET-CT
- ▲ Balneophototherapie bei dermatologischen Indikationen
- ▲ Vakuum-Versiegelungstherapie (V.A.C.) von Wunden in der ambulanten Versorgung
- ▲ Interstitielle Brachytherapie bei lokal begrenztem Prostatakarzinom
- ▲ Nicht medikamentöse, lokale Behandlung der benignen Prostata-Hyperplasie (BHP)

36.2.2 Anwendung evidenzbasierter Grundsätze außerhalb der Verfahrensordnung

Für eine Reihe von Richtlinien und Entscheidungen des G-BA gelten zwar die allgemeinen Grundsätze der Verfahrensordnung, für die eigene Arbeit wurden jedoch weitere, detaillierte Maßstäbe entwickelt. Diese enthalten häufig Grundanforderungen der Evidenzbasierten Medizin.

So ist es die Aufgabe des G-BA, geeignete Präparate in Festbetragsgruppen zusammenzufassen, die einheitlich vergütet werden. Traditionell ist die Bewertung von Arzneimitteln einer der Hauptgegenstände der evidenzbasierten Gesundheitsversorgung. Der gleichnamige Unterausschuss Arzneimittel nach § 91 Abs. 5 des G-BA hat es im Jahr 2004 vorgezogen, eigene Grundsätze für seine Arbeit, nämlich die evidenzbasierte Zuordnung von Arzneimitteln, festzulegen. Dieses Vorgehen entspricht einer verwaltungsinternen Vorschrift, die über ihre regelmäßige Anwendung eine Bindungswirkung entfaltet.

Durch die detaillierten gesetzlichen Vorgaben im SGB V bedingt arbeitet der Unterausschuss Disease-Management-Programme nach § 92 Abs. 5 SGB V auf der Grundlage von Leitlinien, die nach evidenzbasierten Gesichtspunkten erstellt wurden. Auch dieser Unterausschuss gibt sich derzeit (Ende 2006) eigene Grundsätze, die der Strukturierung der eigenen Arbeit dienen. Perspektivisch möglich erscheint, dass beide Unterausschüsse nach Erprobung ihrer Arbeitsgrundsätze im Alltag diese in die Verfahrensordnung des G-BA einbringen.

36.3 Erfahrungen mit der Anwendung der Prinzipien der EBHC im G-BA

Es ist der umfassenden Aufbereitung der vorhandenen Wissens- und Informationsbasis

geschuldet, dass die Begründung für Entscheidungen aufwendiger geworden ist. Die Einführung neuer Leistungen aufgrund weniger Experteneinschätzungen ist zum seltenen Einzelfall geworden, was gleichzeitig Vor- und Nachteile in sich birgt. Während die Befragung von Fachexperten zu einer relativ raschen Entscheidungsfindung führen kann, ist eine Beratungsdauern von bis zu zwei Jahren bei Anwendung evidenzbasierter Grundsätze keine Seltenheit.

Auch die Etablierung des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) durch den Gesetzgeber hat hierfür keine kurzfristige Besserung erkennen lassen. Häufig wird mit der eingehenden Prüfung und der damit verbundenen Verfahrensdauer der Vorwurf der Rationierung und der Behinderung von Innovationen in Zusammenhang gebracht.

Dem kann entgegengestellt werden, dass die Verwendung von Versicherungsgeldern in einem speziellen Bereich bei Vorgabe eines festen Budgets immer bedeutet, dass dieses Geld an anderer Stelle fehlt. Eine bewusste und informierte Entscheidungsfindung ist deshalb eine *conditio sine qua non* für die Organisation des Gesundheitssystems. Die Methoden der Evidenzbasierten Medizin sind hierbei eine wertvolle Hilfestellung und Grundlage.

Nur schwer lösbar erscheint zudem der Widerspruch zwischen allgemeinverbindlichen Richtlinien für alle an der Gesundheitsversorgung Beteiligten und den spezifischen und oft Einzelfälle darstellenden Versorgungssituationen im Alltag. Beiden Aspekten gerecht zu werden, der einheitlichen Regulierung des Gesundheitssystems über Richtlinien einerseits und der individuellen Gesundheitsversorgung des Einzelnen (unter anderem durch begründetes Abweichen von einer Richtlinie im Einzelfall, entsprechend dem Konzept der EbM) andererseits, kann einer Quadratur des Kreises gleichkommen. Dieses dem Lipski-Dilemma [11] entspre-

chende Problem ist jedoch weniger der EbM als den gesetzgeberisch vorgesehenen Entscheidungsoptionen des G-BA geschuldet. Der Gesetzgeber hat bei der Anwendung von Arzneimitteln im ambulanten Bereich erstmals die Möglichkeit zugelassen, dass Ärzte in begründeten Ausnahmen von den Richtlinien des G-BA zur Anwendung von Arzneimitteln abweichen können.

Essenziell ist das richtige Gleichgewicht zwischen Grundsatzregelungen auf dem Wege von Richtlinien und dem erforderlichen Entscheidungsspielraum auf der individuellen Ebene der Patientenbetreuung. Das hieraus resultierende Maß an Entscheidungsfreiheit darf dann allerdings nicht als therapeutische und diagnostische Beliebigkeit verstanden werden.

Bei der Anwendung der zuvor gültigen Verfahrensrichtlinie zeigte sich, dass ein vollständig sektorunterschiedliches Vorgehen bei der Bewertung derselben Untersuchungs- und Behandlungsmethode zumeist nicht sinnvoll erscheint. Am Beispiel der Positronenemissions-Tomographie wurde deutlich, dass die Evidenzbewertung auch für beide Sektoren gemeinsam vorgenommen werden kann. Die sektorspezifische Betrachtung der Notwendigkeit und Wirtschaftlichkeit erfolgt danach wieder getrennt. Auch besteht grundsätzlich die Möglichkeit, einheitliche Qualitätsanforderungen zu formulieren, sofern dies im Sinne einer sachgerechten Versorgung sinnvoll erscheint. Für Patienten ist der Ort der Leistungserbringung grundsätzlich sekundär, sodass bei vergleichbarem zu behandelndem Patientenspektrum im Hinblick auf die Erbringung der Leistung ein einheitlich hohes Qualitätsniveau erwartet werden kann.

Die Voraussetzungen für eine gemeinsame Evidenzbewertung hat der G-BA mit seiner neuen Verfahrensordnung durch die Einführung zahlreicher sektorübergreifender Themengruppen bereits geschaffen. Über eine mögliche Angleichung der unterschiedlichen gesetzlichen Regelungen zur Metho-

denbewertung im ambulanten und stationären Sektor bestehen verständlicherweise unterschiedliche Positionen bei den einzelnen Vertragspartnern im G-BA. Während die eine Seite eine grundsätzliche Zulassung von Innovationen auch für den stationären Bereich wünscht, sieht die andere Seite in einer solchen Entwicklung eine erhebliche Bedrohung für den Innovationsstandort Deutschland. Insofern stellt die in der neuen Verfahrensordnung geschaffene Einrichtung von zahlreichen themenübergreifenden Arbeitsgruppen, wie im Falle der Positionenemissions-Tomographie, und die Öffnung der Ausschüsse für alle Bänke einen Kompromiss zur Begegnung der Probleme dar.

Nicht selten ist die „Überstrapazierung der EbM“ als Prozedur zur Erreichung allenfalls mittelfristiger gesundheitspolitischer Ziele einzelner Interessensgruppen zu beobachten. Hierzu gehören u.a. folgende Muster:

1. Janusköpfigkeit: Selektive Nutzung der EbM in vorteilhaften Situationen:
 - „Forderung hoher Evidenzklassen bei der Einführung von Mindestmengen“ versus „Forderung niedrigerer Evidenzklassen bei der Einführung von Innovationen“.
 - „Forderung, dass GKV-Medizin evidenzbasiert sein muss“ versus „Ignorieren der Strategien der EbM bei Maßnahmen der Wellness- und Komfortmedizin oder bei so genannten Individuellen Gesundheitsleistung (IGeL)“.
 - „Niedrige Evidenz bei (extrabudgetär finanzierten) Maßnahmen der Prävention“ versus „hohe Evidenz bei der (aus dem Budget finanzierten) Kuration“.
2. EbM als „Hürde“ mit eigenen Grundsätzen und Begrifflichkeiten, die es Außenstehenden unmöglich macht, an der Diskussion zu partizipieren (technokratisches Allokationsinstrument).
3. EbM als „Entschuldigung“ in der Versorgungsplanung: „Wir würden diese oder jene Methode ja gern vergüten, können

- dies aber nicht aufgrund fehlender Evidenzbasierung“ (Externalisierung des Problems).
4. Ablehnen von Technologien, die unzureichend beforscht sind, ohne sich ernsthaft um die desolante Primärstudien-situation in Deutschland zu kümmern. EbM kann nur dann funktionieren, wenn genügend hochqualitative Primärstudien gefördert und durchgeführt werden.
 5. EbM als „Gütezeichen“: Die Entscheidung muss gut sein, da sie evidenzbasiert getroffen wurde. EbM wird als inhaltsleere Hülse missbraucht.

Entscheidungen des G-BA werden regelmäßig insbesondere vor Sozialgerichten auf den Prüfstand gestellt. Hierbei hat sich gezeigt, dass solche Richtlinien, die auf der Grundlage von HTA-Berichten erstellt wurden, einer Überprüfung standhalten.

Besonders stellt sich die Situation im Arzneimittelbereich dar: Gerade die Zusammenführung von Präparaten in Produktgruppen wird regelmäßig durch die Pharmaindustrie beklagt. Auch hier zeigte sich bislang, dass die Arbeitsweise des G-BA nicht revidiert werden musste.

Unbenommen hiervon ist der nach wie vor diskutierte Legitimationsrahmen des G-BA. In einer Entscheidung des Bundesverfassungsgerichts aus dem Jahre 2006 wurde dem G-BA nicht das Recht zugebilligt, in therapeutisch schwierigen Situationen Einschränkungen des Leistungskatalogs vorzunehmen. Die Auswirkungen des Urteils auf die Richtlinienggebung des G-BA werden derzeit diskutiert.

36.4 Ausblick

Zusammenfassend lässt sich feststellen, dass mit der Einführung evidenzbasierter Grundsätze in den Verfahrensablauf des G-BA bei allen Problemen ein rechtssicheres und ver-

lässliches Verfahren etabliert werden konnte. Die zunächst im ambulanten Bereich begonnene Einführung wurde vom stationären Sektor übernommen, was internationalen Vorgehensweisen entspricht. Eine Rückkehr zu einer ausschließlich expertenbestimmten Verfahrensweise erscheint deshalb derzeit ausgeschlossen.

Noch nicht suffizient gelöst scheint die sektorübergreifende Zusammenarbeit in den dafür vorgesehenen Ausschüssen des G-BA (z.B. Themengruppen). Wenngleich die fachliche Bewertung der verfügbaren Evidenz zu meist durch das IQWiG erfolgt, so ergeben sich dennoch im Ausschuss zahlreiche Fragestellungen, wie beispielsweise die Berücksichtigung spezifischer Patientensubgruppen und die sektorspezifische Notwendigkeit. Auch bestehen abweichende Positionen bezüglich der Einführung von neuen Methoden unter Evaluationsbedingungen, die Konzentration des Angebotes auf zertifizierte Einrichtungen oder aber die Einführung unter Erhebung von Zuzahlungen zwischen den einzelnen im G-BA beteiligten Parteien. Gleichwohl handelt es sich hier um Optionen, die in anderen Gesundheitssystemen längst etabliert wurden [12]. Ein Mittelweg zwischen weitgehend unregulierter Neueinführung von Innovationen in der stationären Versorgung (Verbotsvorbehalt) und Zugangsbarrieren für die ambulante Versorgung (Erlaubnisvorbehalt) steht somit noch nicht in Aussicht.

Noch weniger berührt von den Methoden der Evidenzbasierten Medizin ist der Bereich der Qualitätssicherung und des Qualitätsmanagements. Sowohl für die Erstellung von Qualitätsindikatoren als auch für den Nachweis der Wirksamkeit von Maßnahmen zur Qualitätsverbesserung sind die Grundsätze der Evidenzbasierten Medizin anwendbar. Die künftige Weiterentwicklung der Verfahrensordnung im G-BA wird deshalb auch diesen Bereich berücksichtigen müssen. Gleiches trifft auch auf die Prioritätensetzung zu. Nach wie vor arbeitet der G-BA eher reaktiv,

indem Versorgungsprobleme über die Bänke oder das Ministerium an den G-BA herangebracht und durch Richtlinienbeschlüsse gelöst werden. Eine systematische, sektorenübergreifende Prioritätensetzung, die einem geordneten, zielorientierten Vorgehen im Sinne der EBHC entspricht, steht noch aus und wird derzeit in Ansätzen in einzelnen Unterausschüssen entwickelt.

Durch die Einführung der evidenzbasierten Grundsätze wurde zwar ein höheres Maß an Transparenz der Entscheidungsfindung etabliert, die Legitimation des Gremiums wird jedoch weiterhin kritisch hinterfragt. Es entspräche den Prinzipien der EBHC, die Arbeitsweise des G-BA noch weiter zu öffnen. Hierzu gehört die Erweiterung der Antragstellung für die Öffentlichkeit einschließlich der Industrie sowie die Möglichkeit eines Widerspruchsverfahrens, wie es beispielsweise NICE in Großbritannien etabliert hat. Welche Rolle die Entscheidungsfindung durch Repräsentativorgane wie die durch die Bänke besetzten Bundesausschüsse künftig haben wird, ist dabei mit Blick auf das Anfang 2007 durch den Bundestag verabschiedete Gesetz zur Stärkung des Wettbewerbs offen. Der holländische Ziekenfondsraad wurde schon vor 6 Jahren in ein Gremium (College voor Zorgverzekeringen – CVZ) umgewandelt, das mit vom Ministerium benannten Fachexperten besetzt wird. Die Entscheidungen werden dort nicht mehr durch die gewählten Vertreter der einzelnen Sektoren getroffen. Ob diese in Holland offensichtlich erfolgreiche Vorgehensweise in Deutschland die Popularität der evidenzbasierten Entscheidungsfindung fördern wird, darf hinterfragt werden.

Literatur

[1] Schimmelpfeng-Schutte R, Richtliniengestaltung durch den Bundesausschuss der Ärzte und Krankenkassen und demokratische Legitimation. *Neue Z Sozialrecht.* (1999), 11, 530–537

- [2] Rawls J (1999) *A theory of justice*. Revised edition. Harvard Univ. Pr., Cambridge
- [3] Leventhal GS (1980) What should be done with equity theory? New approaches to the study of fairness in social relationships. In: Gergen K, Greenberg M, Willis R, *Social exchange: Advances in theory and research*, 27–55. Plenum Pr., New York
- [4] Gibis B (2001) *Expertengremien in der Politikberatung*. In: Fozouni B, Güntert B, *Prioritätensetzung im deutschen Gesundheitswesen*, 164–173. Logos, Berlin
- [5] Muir Gray JA (2001) *Evidence-based Healthcare. How to make health policy and management decisions*. Churchill Livingstone, Edinburgh
- [6] Wild C, Gibis B, *Evaluations of health interventions in social insurance-based countries: Germany, the Netherlands, and Austria*. *Health Policy.* (2003), 63, 187–196
- [7] Jung K et al., *Bundesausschuss der Ärzte und Krankenkassen: Ansprüche der Versicherten präzisieren*. *Dtsch Arztebl.* (2000), 97, A-365–370
- [8] Gibis B, Schmacke N, Windeler J, *Akupunktur. Ergebnisse und Zweifel. Modellversuche der GKV zur Akupunktur ermöglichen gezielte wissenschaftliche Erkenntnisse zu einer Therapiemethode, über deren spezifische Wirksamkeit Zweifel bestehen*. *Dtsch Arztebl.* (2001), 98, A-445–447
- [9] Gibis B, Rheinberger P, *Erfahrungen mit und Impact von Health Technology Assessment im Bundesausschuss der Ärzte und Krankenkassen*. *Z Arztl Fortbild Qualitätssich.* (2002), 96, 82–90
- [10] *Gemeinsamer Bundesausschuss (G-BA), Verfahrensordnung des Gemeinsamen Bundesausschusses*. *Bundesanzeiger.* (2005), 16, 998
- [11] Schwartz FW, Wismar M, Amelung V (2003) *Planung und Management*. In: Schwartz FW et al., *Das Public Health Buch. Gesundheit und Gesundheitswesen*, 695–713. Urban und Fischer, München
- [12] Gibis B, Koch-Wulkan P, Bultman J (2004) *Shifting criteria for benefit decisions in social health insurance systems*. In: Saltman R, Busse R, Figueras J, *Social health insurance systems in western Europe*, 189–206. Berkshire, Open Univ. Pr.

37 Evidenzbasierte Medizin und ethische Aspekte

Nikola Biller-Andorno, Christian Lenk

Das Grundanliegen der Evidenzbasierten Medizin ist zugleich auch ein moralisches: Es geht darum, bei Entscheidungen bezüglich der Versorgung individueller Patienten die gegenwärtig beste Evidenz einzusetzen, und zwar in gewissenhafter, expliziter und abgewogener Weise [1]. In die Sprache ethischer Prinzipien übersetzt heißt dies: EbM soll helfen, **Schaden zu vermeiden** und **Nutzen zu maximieren**, und zwar zunächst auf der individuellen Ebene. Dabei ist diese Definition von EbM durchaus mit einem weiteren wichtigen Prinzip der Medizinethik vereinbar, dem **Respekt vor der Patientenautonomie**. Denn sie fordert keine ausschließliche, sklavische Orientierung an der gegenwärtig besten Evidenz als einzigem Kriterium, sondern lässt durchaus Spielraum für andere Erwägungen, die im Einzelfall hinzukommen mögen, wie besondere Präferenzen oder Abneigungen des Patienten.

Ein genauerer Blick auf die wissenschaftliche Datenlage im Hinblick darauf, ob eine Intervention mehr schadet als sie nützt und welche von mehreren die beste ist, zieht zugleich zwei weitere Fragen nach sich: erstens, ob sich manche Interventionen angesichts geringen Nutzens überhaupt rechtfertigen lassen, und zweitens, welcher von ungefähr gleich wirksamen Interventionen der Vorzug zu geben sei. An dieser Stelle liegt es nahe, die individuelle Ebene zu verlassen und die Erwägungen mit gesundheitsökonomischen Gesichtspunkten zu koppeln [2–4]: „Rechnet“ sich eine Intervention angesichts ihres belegbaren Nutzens, auch im Vergleich zu konkurrierenden Produkten oder Verfahren? Nun ist auch diese Frage per se nicht amoralisch. Im

Gegenteil zielt sie auf eine rechtfertigbare, im besten Falle **gerechte Verteilung** begrenzter Ressourcen im Gesundheitswesen.

Nichtsdestotrotz birgt der Transfer der EbM, die zunächst als medizinische Entscheidungshilfe zum Wohle des individuellen Patienten konzipiert war, auf die Makroebene Risiken, die es zu berücksichtigen gilt. Zudem ist zu fragen, inwiefern der „Trend“ zur EbM die medizinische Forschungsagenda beeinflusst und möglicherweise verengt. Diese drei Anwendungskontexte der EbM – medizinische Forschung, Ressourcenallokation im Gesundheitswesen und Verbesserung der Versorgungsqualität – sollen im Folgenden genauer beleuchtet werden. Die Ausführungen stützen sich hierbei auf die Ergebnisse einer von der Europäischen Kommission geförderten Projektgruppe (EVIBASE), die sich in den Jahren 2000 bis 2003 mit „Ethical Issues of Evidence-based Practice in Medicine & Health Care“ befasst hat [5; 6].

37.1 EbM in der Forschung

Die epidemiologische und statistische Stärke der EbM beruht auf ihrer Präferenz für die Durchführung von randomisierten kontrollierten Studien (Randomized Controlled Trials, RCTs). Dadurch entsteht eine Hierarchie des Wissens, an deren unterem Ende die alterhergebrachte Meinung der Autoritäten und an deren oberem Ende gesichertes experimentelles Wissen steht [7]. Diese Hierarchie bevorzugt allerdings aus experimentell-methodischen Gründen die „typische“ Medikamentengabe vor komplexeren klinischen

bzw. therapeutischen Interventionen. Es ist nicht klar, wie Eingriffe, die nicht in RCTs getestet werden können, in der bestehenden EbM-Hierarchie eingeordnet werden können. Im Zweifelsfall geraten sie jedoch in Verdacht, dass es keine oder nur eine ungenügende Evidenz für ihre Durchführung gibt. Das führt dazu, dass Randbereiche der klinischen Medizin bzw. Schulmedizin, wie z.B. die Psychotherapie (bei der, wenn man so will, der Therapeut selbst zugleich einen Teil der Therapie ausmacht), unter verstärkten Legitimationsdruck hinsichtlich der Wirksamkeit ihrer Behandlung geraten [8].

In RCTs werden Patienten meist hochselektiv für eine spezifische Behandlung ausgewählt. Es ist daher naheliegend, dass die Studienergebnisse besser sind als die Behandlungsergebnisse in der Praxis sein können. So konnten Heiat et al. [9] im US-amerikanischen Setting zeigen, dass Patienten mit Herzversagen in RCTs sich hinsichtlich ihrer demographischen Daten signifikant von der Gesamtgruppe der Patienten mit Herzversagen unterscheiden. Die Studienpatienten waren tendenziell jünger, häufiger männlich, hatten bestimmte klinische Merkmale und gehörten eher der angloamerikanischen

Tab. 37.1: Empfehlungen der EVIBASE-Gruppe zu EbM und Forschungsprioritäten [6]

- Die EbM-Bewegung ist wegen des engen Spektrums akzeptierter methodologischer Herangehensweisen kritisiert worden, besonders hinsichtlich der dominanten Stellung randomisierter kontrollierter Studien (RCTs). Die EbM-Bewegung hat auf diese Kritik reagiert, indem sie versucht hat, die Bandbreite akzeptierter Evidenz um Nicht-RCT-Daten zu erweitern. Diese Anstrengungen sind vielversprechend und sollten weiter verfolgt werden. Sie werden jedoch mit dem kritischen Geist und den sorgfältigen Methoden unternommen werden müssen, die für EbM charakteristisch sind.
- Die gesamte Methodologie medizinischer Forschung betont insbesondere die randomisierte Studie, die für die Untersuchung einfacher Interventionen (üblicherweise der Verabreichung von Medikamenten) entwickelt worden ist. Es gibt jedoch mehr und mehr Hinweise darauf, dass viele komplexe Interventionen (Prävention, die Bedeutsamkeit persönlicher Zuwendung im Behandlungskontext) mindestens so wirksam sind wie Medikamente. Die Wirksamkeit solcher komplexer Interventionen mittels RCTs zu erweisen ist jedoch schwierig und zudem konzeptionell fragwürdig. Daher sollte in flexiblere Forschungsinstrumente innerhalb der EbM-Bewegung investiert werden.
- Vulnerable soziale Gruppen (wie Alte oder Arme) und „schwache“ Bereiche der Medizin und Gesundheitsversorgung (wie Gesundheitsförderung oder die Versorgung chronisch Kranker) sind in der gegenwärtigen Evidenzbasis immer noch unterrepräsentiert, obwohl Anstrengungen unternommen worden sind, diese Lücke zu schließen. Um die praktische Relevanz von EbM zu stärken, würde Forschung hinsichtlich dieser Gruppen und Bereiche besondere Priorität verdienen. Öffentliche Forschungseinrichtungen haben eine besondere Verpflichtung in dieser Hinsicht.
- Gemessen an ihrer Relevanz sind soziale und kulturelle Faktoren für Gesundheit und für die Wirksamkeit von Gesundheitsversorgung immer noch ein zu wenig beforschtes Gebiet. Idealerweise sollte jeder Review Hinweise auf die verfügbare Evidenz zu den jeweiligen sozialen und kulturellen Faktoren einschließen.
- Um die praktische Bedeutsamkeit der Evidenz zu verbessern, sollten Patientenvertreter und Gesundheitsversorger in die Lage versetzt werden, am Entscheidungsprozess hinsichtlich Forschungsprioritäten teilzunehmen. Dies sollte die Auswahl der Themen, der zu untersuchenden Gruppen und Interventionen einschließen.
- Der Ruf nach mehr evidenzbasierten Therapien heißt nicht, dass mehr Behandlung immer besser ist. In manchen Krankheitsphasen kann es besser sein, weitere kurative Behandlungsversuche zu unterlassen. In dieser Hinsicht könnte EbM, wenn sie über Behandlungserfolge oder -misserfolge informiert, als Verbündeter gegen Überbehandlung oder Medikalisierung wirksam werden.

Bevölkerung an [9; 10]. Dies läuft allerdings dem Anspruch zuwider, mit der Methodik der RCTs zu definitiven Aussagen zur klinischen Wirksamkeit zu kommen. Dazu wäre es nötig, dass die betroffene Patientengruppe annäherungsweise der in die Studie eingeschlossenen Gruppe ähnelt. Eine Optimierung der Behandlung – eines der expliziten Ziele der EbM – kann nur dann erfolgen, wenn die Evidenzbasis auch für die betreffende Personengruppe relevant ist.

37.2 EbM bei der Ressourcenallokation

Von vielen Akteuren im Gesundheitswesen (z.B. Krankenkassen, Kliniken, Gesundheitspolitik, Sozialgerichten, (vgl. [11]) werden die Ergebnisse evidenzbasierter Studien bzw. evidenzbasierter Leitlinien als Grundlage für Entscheidungen zur Ressourcenallokation genommen. Im Hintergrund steht der Gedanke, dass nur wirksame Behandlungen von der öffentlichen Gesundheitsversorgung übernommen werden und die wirksamere vor der unwirksameren Behandlung bevorzugt werden sollte. Fraglich ist jedoch, wie sich diese Konstellation auf bestimmte Patientengruppen wie etwa chronisch Kranke

oder Behinderte auswirkt, deren Grunderkrankungen zumeist dadurch charakterisiert werden, dass es keine wirksame Therapie gibt. Eine verstärkte Orientierung an Kriterien klinischer Wirksamkeit würde daher dazu führen, dass solche Patientengruppen weniger gut versorgt werden als bisher. Aus diesem Grunde wäre es zur Erhöhung der Akzeptanz evidenzbasierter Untersuchungen sinnvoll, neben die Fokussierung auf die Methode der RCTs eine stärkere Einbeziehung der betroffenen Patientengruppen zu stellen.

Weiterhin erscheint der Trend problematisch, die Feststellung der Evidenz in staatlichen oder halbstaatlichen Institutionen wie z.B. NICE in Großbritannien (vgl. [12]) zu monopolisieren. Die dezentrale Arbeit verschiedener Forschungsgruppen mit größerer Unabhängigkeit von der Gesundheitspolitik scheint im Interesse der Patienten angemessener zu sein.

37.3 EbM und Versorgungsqualität

Wie Praetorius und Sahn [13] dargelegt haben, sollte die Patientenautonomie, das Arzt-Patient-Verhältnis, aber auch die Therapiefreiheit des Arztes durch eine umfassende „Leitlinisierung“ der Medizin nicht zu sehr

Tab. 37.2: Empfehlungen der EVIBASE-Gruppe zu EbM und Allokationsfragen [6]

- Es besteht Konsens, dass das Konzept der Gerechtigkeit einer der Eckpunkte eines ethischen Rahmens zur EbM ist. Die Qualität der Gesundheitsversorgung für vulnerable Gruppen ist ein wichtiger Maßstab, wenn es darum geht, die Folgen der EbM in ethischer Hinsicht zu bewerten. EbM sollte vulnerable soziale Gruppen in die Forschung ebenso einschließen wie „schwächere“ Disziplinen innerhalb der Medizin, auch wenn es schwierig oder unmöglich ist, in diesen Bereichen RCTs durchzuführen.
- Bislang sind methodische Instrumente zur Einschätzung der Wirksamkeit und des Nutzens von Interventionen vornehmlich auf Vergleiche innerhalb bestimmter Subdomänen in der Medizin geeicht. Um die Relevanz von EbM für Allokationsentscheidungen zu verbessern, wäre es wünschenswert, mehr und bessere Instrumente für den Vergleich zwischen den Subdomänen zu entwickeln.
- Um die Legitimität evidenzbasierter Allokationsverfahren zu stärken, ist maximale Transparenz hinsichtlich der angewandten Verfahren unabdingbar. Transparenz in diesem Sinne könnte durch die Existenz einer Pluralität an Institutionen gefördert werden, die in den Prozess der Einschätzung von Evidenz für Allokationsfragen involviert sind. Grundsätzlich ist es nicht erstrebenswert, dass eine einzige Institution ein Monopol in diesem Prozess hat.

belastet werden. Es scheint auch nicht in der EbM als solcher angelegt zu sein, dass sie teilweise von den Verwaltungen und Bürokratien in Form eines repressiven Druckmittels angewandt wird. Stattdessen sollte man EbM als ein Instrument verstehen, behandelnden Ärzten mehr und bessere Therapieoptionen zu erschließen. In diesem Sinne stellt sich die EbM auch in der Mediziner Ausbildung als ein nützliches Instrument dar, um den Studierenden einen Weg an die Hand zu geben, systematisch nach wissenschaftlichen Behandlungsmöglichkeiten für ihre Patienten zu suchen.

Problematisch bei der Produktion von Evidenz erscheint allerdings, dass die geforderten Fallzahlen und Studiengrößen oft genug nur unter Beteiligung der Industrie erreicht werden, während andere therapeutische Möglichkeiten, die aus Sicht des Kliniklers oder aufgrund realer Versorgungsdefizite wichtig sind, aufgrund mangelnder Finanzierung weniger gut untersucht werden können. Solche Ungleichheiten führen not-

wendigerweise zu einem Bias bei der Beurteilung, welche Medikamente für die medizinische Versorgung am besten geeignet sind. Zu ihrer Beseitigung dürfte das Engagement der Gesundheits- und Forschungspolitik notwendig sein.

37.4 Schlussfolgerungen

- ▲ Ergebnisse evidenzbasierter Studien müssen im Zusammenhang mit der angewandten Methodik und den untersuchten Patientengruppen beurteilt werden.
- ▲ Mangel an Evidenz aus Studien bedeutet nicht zwangsläufig mangelnde therapeutische Wirksamkeit.
- ▲ Vorschnelle Schlüsse von Evidenz auf ökonomische Effizienz sollten vermieden werden.
- ▲ Das Potenzial der EbM kann am besten genutzt werden, wenn Raum bleibt für begründete Präferenzen von Arzt/Ärztin und Patient.

Tab. 37.3: Empfehlungen der EVIBASE-Gruppe zu EbM und Verbesserung der Versorgungsqualität [6]

- Neben dem Konzept der Gerechtigkeit gehört der Respekt vor der Autonomie des Patienten zum ethischen Rahmenwerk für die Beurteilung der Auswirkungen von EbM. Um den Respekt vor der Autonomie hochzuhalten und zu bewahren, ist der gemeinsame Erwägungsprozess von Patient und Arzt bezüglich Behandlungsoptionen zentral. EbM sollte als ein Instrument betrachtet werden, das diesen Prozess stärkt und nicht die Optionen auf eine einzige „Rationale“ reduziert, wobei alle anderen Optionen zu „Irrationalen“ stilisiert werden.
- EbM muss als Instrument für das Empowerment, nicht die Einengung von Ärzten und den Angehörigen anderer Heilberufe verstanden werden. Der ganze Prozess des Übersetzens der Evidenz in einen praktischen Anwendungskontext muss in einer Weise geformt werden, die Medizinern Nutzen verspricht für die praktischen Probleme und Entscheidungen, mit denen sie sich konfrontiert sehen.
- Es ist wichtig, dass alle zukünftigen Ärzte (und, soweit angemessen, auch Angehörige anderer Heilberufe) lernen, die Grundprinzipien und -verfahren der EbM wie auch den potenziellen praktischen Nutzen der EbM zu verstehen. Das zentrale Ziel solcher Anstrengungen in der Lehre sollte sein, eine (selbst-)kritische Einstellung gegenüber allen Ansprüchen auf Evidenz zu fördern.
- EbM ist so wichtig, dass sie nicht auf Ärzte beschränkt sein sollte. Eine Konsequenz aus dieser Feststellung ist, dass EbM-Prinzipien und -Verfahren – unter Berücksichtigung der erwähnten Caveats – auch auf andere Felder im Gesundheitswesen und der Gesundheitsförderung ausgedehnt werden sollten. Eine weitere Konsequenz ist, dass ein integrierter Ansatz hinsichtlich der Produktion und Implementierung von Evidenz wünschenswert wäre. So sollten zum Beispiel die Angehörigen anderer Heilberufe in den Prozess der Erarbeitung von Daten und die Erstellung von medizinischer Leitlinien einbezogen werden.

Literatur

- [1] Sackett DL et al., Evidence based medicine: what it is and what it isn't. *BMJ*. (1996), 312, 71–72
- [2] Biller-Andorno N, Evidenz-basierte Gesundheitsversorgung – kritische Nachfragen vor medizinethischem Hintergrund. *Wien Med Wochenschr.* (2002), 152, 161–164
- [3] Biller-Andorno N (2002) Medizinethische Implikationen einer evidenz-basierten Gesundheitsversorgung. In: Brand A et al., *Individuelle Gesundheit versus Public Health*, 113–125. LIT-Verl., Münster
- [4] Hope T, Evidence based medicine and ethics. *J Med Ethics.* (1995), 21, 259–260
- [5] ter Meulen R et al. (2005) Evidence-based practice in medicine & health care. A discussion of the ethical issues. Springer, Berlin
- [6] Houtepen R, Biller-Andorno N, Dickenson D, Vineis P, Lie R, ter Meulen R (2005) Recommendations from the Evibase Project. In: ter Meulen R et al., *Ethical issues of evidence-based practice in medicine and health care*, 173–176. Springer, Berlin
- [7] Oxford Centre for Evidence Based Medicine (CEBM). Levels of evidence and grades of recommendation. http://www.cebm.net/levels_of_evidence.asp (04.12.06)
- [8] Henningsen P, Rudolf G, Zur Bedeutung der Evidence-Based Medicine für die Psychotherapeutische Medizin. *Psychother Psychosom Med Psychol.* (2000), 50, 366–375
- [9] Heiat A, Gross CP, Krumholz HM, Representation of the elderly, women, and minorities in heart failure clinical trials. *Arch Intern Med.* (2002), 162, 1682–1688
- [10] Rogers W (2005) Evidence-based medicine and equity: The exclusion of disadvantaged groups. In: ter Meulen R et al., *Ethical issues of evidence-based practice in medicine and health care*, 129–138. Springer, Berlin
- [11] Biller-Andorno N, Lenk C, Ethische Aspekte der Evidenz-basierten klinischen Medizin. *Z Arztl Fortbild Qualitatssich.* (2003), 97, 681–687
- [12] Harris J, It's not NICE to discriminate. *J Med Ethics.* (2005), 31, 373–375
- [13] Praetorius F, Sahn S, Ethische Aspekte der Regularisierung ärztlichen Handelns. *Ethik Med.* (2001), 13, 221–242

38a Evidenzbasierte Medizin und Recht: Haftungsrecht

Dieter Hart

38a.1 Einleitung

Die Feststellung eines Behandlungsfehlers in einem Arzthaftungsprozess ist ohne das Gutachten eines gerichtlichen Sachverständigen nicht möglich. Die Qualität der gerichtlichen Entscheidung und die des Gutachtens sind abhängig von der Bereitschaft und Fähigkeit des Gutachters, die vorhandene medizinische wissenschaftliche Evidenz, die ärztliche Erfahrung und die Akzeptanz der Behandlung in der Profession zu erkennen, zur Kenntnis zu nehmen, zu bewerten und auf den Einzelfall anzuwenden. Fehlt es an umfassender Kenntnis, Bereitschaft und/oder Fähigkeit, ist das gutachterliche Urteil willkürgefährdet.

Wenn also das gutachterliche Urteil über den Behandlungsfehler (mit) entscheidet, das Gutachten aber möglicherweise „willkürlich“ ist, dann ist in der Folge auch die Qualität des gerichtlichen Urteils zweifelhaft. Eine ärztliche Befürchtung gegenüber rechtlicher Willkür muss aber so gesehen und gerechterweise mindestens auch Furcht vor gutachterlicher Willkür sein. Das Urteil kann immer nur so gut sein wie das sachverständige Gutachten.

Die Methode der EbM könnte für das formelle (Prozessrecht) und materielle Arzthaftungsrecht einen erheblichen Rationalitätsgewinn zumal dann bedeuten, wenn sie in Verbindung mit evidenzbasierten Konsensus-Leitlinien der haftungsrechtlichen Entscheidungsfindung zugrunde gelegt wird. Dies zu begründen ist ein Ziel des folgenden Kapitels.

Der Beitrag zeigt insgesamt **Schnittstellen** zwischen EbM und Haftungsrecht auf

und erörtert, unter welchen Bedingungen EbM die haftungsrechtliche Entscheidungsfindung beeinflusst.

38a.2 Haftungsrecht als Verhaltensnormenrecht

Haftungsrecht hat primär eine retrospektive Ausgleichs- und sekundär eine prospektiv präventive, risikovorbeugende Funktion. In beiden Funktionen ist das Haftungsrecht auch Arzt- oder Medizinhaftungsrecht. Einerseits stellt das Haftungsrecht rückblickend eine Pflichtverletzung bei der Behandlung (behandlungs-, aufklärungs-, dokumentations- und organisationsbezogene Pflichtverletzungen) und dabei einen Behandlungsfehler fest, andererseits impliziert eine solche Feststellung immer zugleich die zukunftsorientierte Pflicht, solche Fehler prospektiv zu vermeiden. Zur Ausgleichsfunktion tritt eine Vorsorgefunktion. Haftungsrecht statuiert Anforderungen an das Verhalten der Heilberufe, indem es Verhaltenspflichten feststellt oder festsetzt. Werden diese Verhaltenspflichten in einem funktionsfähigen Risikomanagementsystem beachtet, entfalten sie ihre zukunftsgerichtete haftungsvermeidende Wirkung [1].

Bei dieser Etablierung von Verhaltenspflichten ist das Arzthaftungsrecht aber nicht vollständig autonom. Die Feststellung eines Behandlungsfehlers ist nur auf der Grundlage eines medizinischen Urteils möglich. Die Feststellung einer Aufklärungspflichtverletzung trifft das Haftungsrecht im Wesentlichen autonom; Dokumentations- und Orga-

nisationspflichtverletzungen sind nur aufgrund sachverständiger medizinischer Beratung rechtlich zu beurteilen. Das notwendige medizinische Urteil setzt die Auswertung der vorhandenen Erkenntnisse über die Behandlung und ihre Organisationserfordernisse voraus. Hier liegen die Schnittstellen von EbM und Arzthaftungsrecht.

38a.3 Schnittstellen von EbM und Haftungsrecht

EbM ist eine empirische Methode der Verwissenschaftlichung der Medizin durch eine Qualitätsrang-Festlegung für medizinische Evidenzen. Die Praxis der EbM bedeutet die Integration individueller klinischer Expertise mit der besten externen Evidenz aus systematischer Forschung. EbM sucht also nach der „best available evidence“, ohne allerdings die anderen Arten der Evidenz gering zu schätzen. EbM macht insofern keine „Richtigkeitsvorgaben“ für Behandlungsziele, sondern „Verfahrensvorgaben“ für die Evaluation von medizinischen Erkenntnissen. EbM kann Beiträge zur Qualitätsverbesserung von Behandlungen, zur Verringerung von Über- und Unterversorgung im Gesundheitssystem und zur gesundheitsökonomischen Bewertung leisten. Die Methode trägt durch Transparenzsteigerung erheblich zu einer Rationalisierung des Prozesses der Generierung von Normen guter Behandlung bei. Die möglichen Auswirkungen der Anwendung von EbM im Medizinrecht und dort im (Arzt-)Haftungsrecht sind bisher kaum systematisch behandelt worden [2–4].

Zwischen EbM und Arzthaftungsrecht kann es folgende Schnittstellen geben, wobei mit Schnittstellen gemeint ist, dass EbM bei der Behandlung arzthaftungsrechtlicher Fragen zu berücksichtigen ist bzw. einen Beitrag zu ihrer Beantwortung leisten kann. Bei der Feststellung eines/r

- ▲ Behandlungsstandards (vgl. Kap. 38a.3.1),
 - ▲ Aufklärungspflicht (vgl. Kap. 38a.3.2),
 - ▲ Dokumentationspflicht (vgl. Kap. 38a.3.3) und
 - ▲ Organisationspflicht (vgl. Kap. 38a.3.4)
- sollte EbM in verschiedenen Beratungs- bzw. Entscheidungsfunktionen hilfreich sein.

38a.3.1 Behandlungsstandard

Der **Standard** einer guten Behandlung wird definiert durch wissenschaftliche Erkenntnis, ärztliche Erfahrung und professionellen Konsens. Der medizinische Standard bestimmt den rechtlichen Standard guter Behandlung; beide sind identisch ([5], bes. S. 53–70; [6]). Der Inhalt der vertraglichen Hauptleistungspflicht „Behandlung“ wird durch den Standard festgelegt, es sei denn, es gäbe eine abweichende individuelle Vereinbarung – was prinzipiell möglich ist – und dasselbe gilt auch für den Inhalt der deliktsrechtlichen Verkehrspflicht (Berufspflicht). Die Nichterfüllung der Pflicht „gute Behandlung“ ist die Pflichtverletzung, die zur Haftung führen kann. Die Medizin bestimmt damit auch über den haftungsrechtlich relevanten Behandlungsfehler.

Der medizinische Standard ist also die **Norm**, die die erforderliche Qualität bzw.

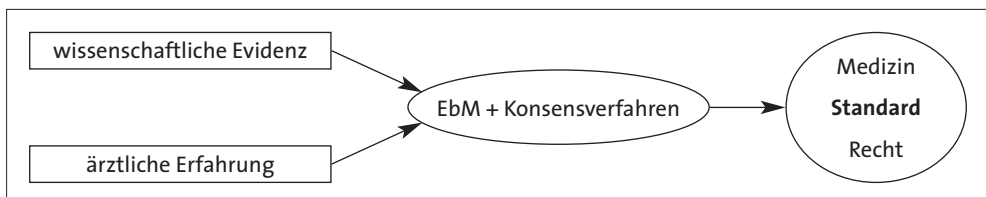


Abb. 38a.1: Medizinischer Standard

den zu erreichenden Nutzen einer medizinischen Behandlung bestimmt, ohne dass damit der individuelle Erfolg einer Behandlung festgelegt werden könnte. Qualität bzw. Nutzen einer Behandlung heißt die abwägende positive Bewertung sowohl ihrer Wirksamkeit wie ihrer Risiken [7]. Es ist allerdings zu beachten und wahrscheinlich ein strukturelles Problem der EbM, dass Studienevidenz in erster Linie Wirksamkeitsevidenz und viel weniger Risikoevidenz ist.

Die **Qualitätsrangskala** der Evidenz (die unterschiedlichen Evidenzlevel) ist eine wichtige **Orientierungshilfe** bei der Feststellung eines Standards guter Behandlung. Je hochwertiger die Evidenz („Evidenzlevel“), je geringer die Zweifel an der Qualität, Validität und Relevanz der Evidenz [8] und je eindeutiger der Empfehlungsgrad, desto höher ist die Anleitungskraft der Evidenz für die Standardentwicklung. Dies gilt aber auch umgekehrt. Je geringer der Qualitätsrang, je höher der Zweifel an der Qualität, Validität und Relevanz der Evidenz und je niedriger der Empfehlungsgrad, desto geringer ist ihre Anleitungskraft, und desto größer ist die Bedeutung der Verfahren der Akzeptanzbildung in der ärztlichen Profession für die Entwicklung des Standards als Norm guter Behandlung.

EbM ist aber nicht nur Orientierungshilfe im gekennzeichneten Sinne, sondern auch **Verpflichtungsmotiv**. Für die Bestimmung eines Standards ist die Methode der EbM heranzuziehen, d.h. es ist nach der besten verfügbaren Evidenz zu suchen, und die ordnungsgemäß gesuchte und gefundene Evidenz ist bei der Beurteilung, ob es einen Standard gibt und gegebenenfalls welchen, zu nutzen. EbM zwingt insofern zur Kommunikation über die Qualität der Evidenz und ist ein wichtiges Element der Qualitätssicherung von Behandlungen. EbM ermöglicht Transparenz.

Die hochwertige Evidenz mit hohem Empfehlungsgrad und geringem Qualitätszweifel definiert den Standard; aber jeder

Standard bedarf als allgemeine und abstrakte patientengruppenbezogene Norm der **Anwendung** auf den einzelnen Behandlungsfall – und das ist (auch) das Feld der internen **klinischen Expertise**. Insofern **integriert die EbM die externe wissenschaftliche Evidenz mit der klinischen Erfahrung** („Verwissenschaftlichung der Medizin“) **auch auf der Ebene des haftungsrechtlichen Standards** („Rationalisierung des Rechts“). Die Feststellung des rechtlichen Standards und seiner Verletzung wird durch EbM transparenter.

Im **gerichtlichen Verfahren** über einen möglichen Behandlungsfehler hat der Sachverständige den medizinischen Standard entsprechend der Methode der EbM festzustellen. EbM qualifiziert gleichzeitig das Gericht zur kompetenten Nachfrage.

Leitsatz

EbM ist eine wichtige Orientierungshilfe bei der Feststellung eines Standards guter Behandlung. EbM ist auch Verpflichtungsmotiv. Für die Bestimmung eines Standards ist die Methode der EbM heranzuziehen. Die Feststellung des rechtlichen Standards (Behandlungspflicht) und seiner Verletzung (Behandlungsfehler) wird durch EbM rationaler und transparenter.

38a.3.2 Aufklärungspflicht

Im Gegensatz zum Behandlungsstandard, den die Medizin bestimmt, wird die Pflicht zur Aufklärung und ihr Umfang durch das Recht, insbesondere das vertragliche und deliktische Haftungsrecht definiert. Die Rolle von EbM bei der Festsetzung von Aufklärungspflichten ist eine **dienende**. Die Qualität der durch ärztliche Aufklärung zur Verfügung zu stellenden **Information**

▲ über die Sicherung der Behandlung (Sicherungs- oder therapeutische Aufklärung),

- ▲ über den Verlauf der Erkrankung mit und ohne Behandlung,
- ▲ über die Risiken der Behandlung und
- ▲ über Behandlungsalternativen (Selbstbestimmungsaufklärung)

kann durch die EbM verbessert werden. Die evidenzbasierte Patienteninformation schafft günstigere Bedingungen der Patientenmitwirkung an der Behandlung (Sicherungsaufklärung), für das Shared decision making und für selbstbestimmte Entscheidungen von Patienten (Selbstbestimmungsaufklärung) [9–11].

38a.3.3 Dokumentationspflicht

Nach neueren Untersuchungen zur Kommunikation zwischen und innerhalb von Sektoren und Institutionen spielen Risikokonstellationen aufgrund der Nichterfüllung von Dokumentationsanforderungen bei Teamarbeit, Schichtwechseln, disziplin- und abteilungsübergreifenden Behandlungen eine erhebliche Rolle [12]. Je höherwertig die Evidenz über das Vorkommen solcher (Beinahe-)Fehlerkonstellationen sein wird, desto eindeutiger und klarer werden die Anforderungen an Dokumentationspflichten in insbesondere Kommunikations- und Kooperationsituationen (Übergabe und Übernahme) ausfallen müssen [13].

38a.3.4 Organisationspflicht

Organisationspflichten (des Praxisleiters, Krankenhausträgers, Klinikleiters) beziehen sich auf die **Organisation der Voraussetzungen** guter

- ▲ Behandlung,
- ▲ Patienteninformation,
- ▲ Dokumentation und Einsichtnahme in die Krankendokumentation.

Organisationspflichten im Bereich der Behandlung beziehen sich insbesondere auf de-

ren Qualität. Qualitäts- und Risikomanagement sind Felder des Organisationspflichtenrechts, in denen das Haftungsrecht partiell durch das GKV-Qualitätssicherungsrecht überlagert wird. Zum Qualitätsmanagement gehört auch die Sicherung der Behandlungsqualität durch Instrumente wie Clinical Pathways, die wiederum an hochwertigen Leitlinien bzw. an Standards zu orientieren sind [14]. Gerade in diesem Felde ist die Bedeutung von EbM hoch, weil EbM die Voraussetzungen für die Bewertung der Qualität solcher Qualitätssicherungsinstrumente zur Verfügung stellt. Insofern sind evidenzbasierte Konsensus-Leitlinien [15] (s. Kap. 35) ein sehr bedeutungsvolles Mittel der Qualitätssicherung und für das qualitätsbezogene Organisationspflichtenrecht ein Maß.

Der Organisationspflichtige hat sicherzustellen, dass innerhalb seiner Institution die Voraussetzungen der standardgemäßen bzw. leitliniengemäßen Behandlung gewährleistet sind. Kann er das nicht, muss er eine nicht-standardgemäße Behandlung verhindern (organisatorische Prävention) [16; 17].

Je hochwertiger zukünftig die Evidenz über Risiko- und Behandlungsfehlerkonstellationen sein wird, desto größer sind die Organisationsanforderungen an die Einrichtung eines funktionsfähigen **Risikomanagements**, dessen Aufgabe darin besteht, Risikovermeidung auch als haftungsrechtliches Ziel zu ermöglichen.

38a.4 Standard, Leitlinien und EbM

Die Methode der EbM und das Verfahren der formalisierten Konsensusfindung gewährleisten zusammen wie kein anderes Verfahren, dass die Voraussetzungen zuverlässiger Standardentwicklung und -feststellung erfüllt werden. Hochwertige Leitlinien sind **das** seriöse Normbildungsverfahren in der Medizin. Diese Leitlinien stellen oder setzen den Standard fest. EbM schafft damit die Er-

kenntnisbasis für ärztliche Leitlinien. Bezogen auf den Standardbegriff füllt EbM die Elemente „wissenschaftliche Erkenntnis und ärztliche Erfahrung“ aus.

38a.4.1 Leitlinien und Haftungsrecht

Haftungsrechtlich gilt prinzipiell der Satz: Nur dem Standard entsprechende und nach ihm aktualisierte Leitlinien können eine medizinische und demzufolge eine rechtliche Verbindlichkeit begründen.

Die Leitlinie ist medizinisch verbindlich, wenn sie dem Standard entspricht und ist rechtlich verbindlich, weil sie dem Standard entspricht.

Ausgangspunkt für jedwede Verbindlichkeitswirkung ist danach der medizinische Standard als Synonym für wissenschaftliche und praktische Lernprozesse in der Medizin sowie in der Folge im Medizin- und Gesundheitsrecht (s. Abb. 38a.2).

„Die evidenzbasierte Konsensus-Leitlinie ist medizinisch verbindlich, weil sie den Standard festsetzt; sie ist deshalb auch rechtlich verbindlich.“ ([18], S. 81 ff, S. 96 ff)

Letztere (konträre) Aussage – es sei darauf hingewiesen, dass eine diese Aussage bestätigende Rechtsprechung noch nicht existiert; zur Rechtsprechung insgesamt [4] – folgt m.E. aus der Tatsache, dass wir über kein besseres Verfahren der medizinischen Normbildung verfügen als die Kombination aus EbM und

Konsensfindung nach einem konsentierten formellen Verfahren. Eine solche aktuelle Leitlinie mit hochgradigen Empfehlungen und geringem Qualitätszweifel ist **standardfestsetzend**. Die Leitlinie ist die rechtlich rezipierte medizinische Norm (s. Abb. 38a.3).

Es spricht einiges dafür, **S3-Leitlinien** der AWMF eine solche Qualität zuzusprechen („Vermutung“), was allerdings voraussetzt, dass die aufgestellten Qualitätskriterien und -verfahren tatsächlich eingehalten wurden – also beispielsweise die Angabe von Evidenzlevel und Empfehlungsgrad, aber auch von Interessenkonflikten – und dass keine zu positive Selbstbewertung vorgenommen wurde.

Die **Abweichung** von der hochwertigen Leitlinie ist deshalb prinzipiell in der haftungsrechtlichen Konsequenz ein Behandlungsfehler – eine Vermutung, die auf zwei Wegen widerlegbar ist: durch die plausible Begründung für eine individuelle Abweichung oder die Begründung der „Überholung“ der Leitlinie aufgrund neuer Erkenntnis oder neuer Bewertung (smaßstäbe) ([18], S. 81 ff, S. 96 ff). Letzteres ist die Folge der Erforderlichkeit der Evaluation und Aktualisierung jeder, auch der hochwertigen Leitlinie.

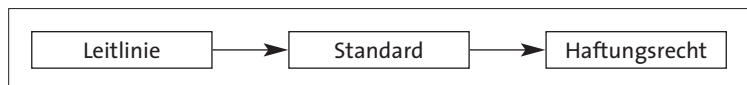
Diese Aussage gilt selbstverständlich nur bei solchen hochwertigen Leitlinien, die einen klar und hochgradig handlungsempfehlenden Charakter haben. Leitlinien, die nur einen informierenden, Behandlungsalternativen aufzeigenden Inhalt haben oder nur einen schwachen Empfehlungsgrad aussprechen oder erheblichen Qualitätszweifeln unterliegen, können diese Funktion nicht übernehmen.

Ein bisher nicht thematisiertes Problem ist die Konsequenz einer „negativen“ Bewertung

Abb. 38a.2: Standardrelevanz



Abb. 38a.3: Leitlinienrelevanz



von hochwertigen Leitlinien durch das (bisher insofern durch den G-BA noch nicht beauftragte) IQWiG. Haftungsrechtlich müsste man in einem solchen Falle wohl von einem möglicherweise fehlenden (z.B. keine ausreichende Evidenz; keine Übereinstimmung bezüglich patientenbezogener Endpunkte), von einem Standard mit Variationsbreite (z.B. kein eindeutiger Konsens bezüglich der Bewertung verschiedener Behandlungsmöglichkeiten) oder mindestens von einem zweifelhaften, unverbindlichen und rechtlich nur indiziellem Befund ausgehen, der die genannte Vermutungswirkung nicht auslöste.

38a.4.2 Leitlinie, Anwendung, Abweichung und Behandlungsfehler

Die **Anwendung** der Leitlinie (Norm) auf den Einzelfall ist (rechtlich) Subsumtion. Im ersten Schritt ist daher festzustellen, ob die Leitlinien-**Indikation** und die Patientenerkrankung übereinstimmen, ob und welche Behandlungsempfehlung abgegeben wird und ob Behandlungsempfehlung und ärztliches Handeln sich entsprechen. Sind diese Voraussetzungen erfüllt, kommt ein Behandlungsfehler nicht in Betracht. Ist die Indikation nicht gegeben, kommt eine Behandlung nach der Leitlinie nicht in Betracht; erfolgt sie dennoch, kann die Leitlinie für die Frage des Behandlungsfehlers nicht herangezogen werden. Es ist nach einer einschlägigen Leitlinie bzw. nach dem entsprechenden Standard zu suchen, falls keine Leitlinie einschlägig ist.

Weicht die Behandlung trotz Indikation von der Leitlinie **ab**, kann der Behandlungsfehleratbestand nur dann vermieden werden, wenn entweder **neue** Erkenntnisse vorliegen oder/und (deshalb) neue Bewertungen vorgenommen werden (müssen) oder wenn **individuelle** patientenbezogene Umstände die empfohlene Behandlung nicht erlauben. Die erste Konstellation ist die Domäne der EbM-bezogenen Aktualisierung. Der Notfall

erfordert, wenn die Leitlinie dafür keine Empfehlung enthält, einen besonderen Standard.

Es ist dem behandelnden Arzt immer zu raten, sowohl die Anwendung der Behandlung nach einer Leitlinie wie die Begründung für die Abweichung von einer Leitlinie zu **dokumentieren**, weil aufgrund fehlerhafter, auch lückenhafter Dokumentation Veränderungen bei der Beweislast in einem Haftungsprozess für die haftungsbegründende Kausalität eintreten können.

Die Ausführungen verdeutlichen, dass die Leitlinie den **Sachverständigen** im Haftpflichtprozess nicht überflüssig macht, aber sie konzentriert seine Aufgabe ([19; 20], bes. S. 121). Innerhalb der Indikation ist die Leitlinie seine „Richtschnur“; er muss sie seinem Auftrag zugrunde legen und anwenden.

Die **Rechtsprechung** der letzten Jahre ist uneinheitlich, gibt den Leitlinien teilweise keine Bedeutung, differenziert teilweise Richtlinien unterschiedlicher Art und Leitlinien nicht voneinander und nimmt insgesamt von der medizinischen Empirie der Leitlinien kaum Kenntnis. Der BGH hat sich bisher mit Leitlinien nicht zu befassen gehabt ([21], bes. S. 107 ff). Es gibt inzwischen allerdings eindeutige richterliche Äußerungen auf einschlägigen Tagungen, die den Gutachter verpflichten, seinem Gutachten vorhandene Leitlinien zugrunde zu legen und sich mit ihnen auseinanderzusetzen.

Leitsatz

EbM und Leitlinien sind die gemeinsame Wurzel guter Behandlungen. Sie sind Instrumente der Standardbeeinflussung und Qualitätssicherung. Hochwertige, aktuelle, evidenzbasierte Konsensus-Leitlinien mit starkem Empfehlungsgrad und geringem Qualitätszweifel setzen den Standard fest und bestimmen damit auch den rechtlichen Standard bzw. umgekehrt den Behandlungsfehler. Solche Leitlinien hat der gerichtliche Gutachter zu beachten. Leitlinien bedürfen der An-

wendung im Einzelfall. Abweichungen sind in der individuellen Behandlung begründet möglich oder sogar geboten. Begründungen für die Abweichung sind zu dokumentieren.

38a.5 Folgerungen für gerichtliche Behandlungsfehlerverfahren

Die Qualität eines Gerichtsurteils über einen möglichen Behandlungsfehler hängt ab von der Qualität des sachverständigen **Gutachtens**. Bisher sind die Gerichte bei der ausdrücklichen Berücksichtigung von ärztlichen Leitlinien für die Standardfeststellung sehr zurückhaltend. Es existiert ein Time Lag hinsichtlich der Wahrnehmung von Leitlinien sowohl bei den Gerichtsgutachtern wie bei den Richtern. Inzwischen scheint diese Phase der Unsicherheit im Umgang mit Leitlinien sich dem Ende zuzuneigen.

Jüngste richterliche Äußerungen verdeutlichen, dass vorhandene hochwertige Leitlinien zukünftig der Arbeit des **gerichtlichen Sachverständigen** zugrunde zu legen sind, d.h. dass sich der Sachverständige mit existierenden einschlägigen Leitlinien bei der Erstellung seines Gutachtens auseinandersetzen hat und dass Gerichte ihm dies auch in den Beweisbeschluss hineinschreiben werden. Insofern ist gleichzeitig und wie bei der Standardfeststellung EbM eine wichtige Grundlage der gerichtlichen Entscheidungsfindung. Der Sachverständige erkennt die Norm guter Behandlung auch mit ihrer Hilfe und wendet den gefundenen Standard bzw. die einschlägige hochwertige Leitlinie auf den Einzelfall an. Die Rationalität des Gutachtens und in der Folge der richterlichen Entscheidung erfährt eine Rationalitätssteigerung. Die Gefahr einer willkürlichen Bewertung wird durch EbM erheblich gesenkt, die Transparenz des Entscheidungsprozesses über den Behandlungsfehler gesteigert.

Die Beachtung der **Dokumentationspflicht** ist auch im Zusammenhang einer leitlinienorientierten Behandlung dringend anzuraten. Fehlt es an der Dokumentation der Abweichungsbegründung und kann der Arzt diese Gründe mit anderen Beweismitteln nicht nachweisen, wird der Behandlungsfehler aufgrund der Abweichung vermutet. Dies folgt aus der Übertragung der ständigen Rechtsprechung zur Verletzung von Dokumentationspflichten ([22], 498 ff, 455 ff). Mangelnde Dokumentation führt wegen der Beeinträchtigung von Nachweismöglichkeiten des Patienten zur den Arzt belastenden Vermutung, die nicht dokumentierte Handlung sei auch nicht erfolgt [23; 24]. Das wiederum kann nach der jüngsten Rechtsprechung des BGH im Falle des groben Behandlungsfehlers zu einer Beweisbelastung des behandelnden Arztes führen, die unterlassene Handlung sei nicht kausal für die Gesundheitsverletzung geworden [25–29].

Leitsatz

Der gerichtliche Sachverständige hat bei der Erstellung seines Gutachtens über einen möglichen Behandlungsfehler vorhandene hochwertige Leitlinien zu berücksichtigen und sich mit ihnen auseinanderzusetzen. EbM ist insofern auch ein Mittel zur Qualitätssteigerung solcher Gutachten und in der Konsequenz des gerichtlichen Urteils über den Behandlungsfehlers.

Literatur

- [1] Becker-Schwarze K, Hart D, Die juristische Sicht: Konflikt zwischen Haftungsrecht und Incident Reporting? GesR. (2005), 4, 1–5
- [2] Hart D, Evidenz-basierte Medizin und Gesundheitsrecht. MedR. (2000), 18, 1–5
- [3] Hart D, Health Technology Assessment (HTA) und gesundheitsrechtliche Regulierung. MedR. (2001), 19, 1–8

- [4] Hart D, Vom Standard zur Leitlinie: Bewertungszusammenhänge im Medizin- und Gesundheitsrecht. Leitlinie und Standard, EbM und HTA, Pflicht und Zweckmäßigkeit. *KritV.* (2005), 88, 154–176
- [5] Hart D (2005) Allgemeiner Teil. In: Hart D, Ärztliche Leitlinien im Medizin- und Gesundheitsrecht. Recht und Empirie professioneller Normenbildung, 23–84. Nomos, Baden-Baden
- [6] Hart D (2005) Leitlinien und Medizinrecht. In: Hart D, Ärztliche Leitlinien im Medizin- und Gesundheitsrecht. Recht und Empirie professioneller Normenbildung, 85–117. Nomos, Baden-Baden
- [7] Hart D, Die Nutzen/Risiko-Abwägung im Arzneimittelrecht. Ein Element des Health Technology Assessment. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* (2005), 48, 204–214
- [8] Falck-Ytter Y, Bericht aus den USA: Eine internationale Arbeitsgruppe entwickelt ein einheitliches System für die Bewertung von Empfehlungen. *Z Arztl Fortbild Qualitätssich.* (2006), 100, 230–231
- [9] Der Patient als Partner im medizinischen Entscheidungsprozess. Förderschwerpunkt des Bundesministeriums für Gesundheit. www.patient-als-partner.de (04.12.06)
- [10] Härter M, Loh A, Spies C (2005) Gemeinsam entscheiden – erfolgreich behandeln. Neue Wege für Ärzte und Patienten im Gesundheitswesen. Dt. Ärzte-Verlag, Köln
- [11] Coulter A, Evidence based patient information is important, so there needs to be a national strategy to ensure it. *BMJ.* (1998), 317, 225–226
- [12] Püschmann H et al., Vollständigkeit und Qualität der ärztlichen Dokumentation in Krankenakten: Untersuchung zu Krankenunterlagen aus Chirurgie, Orthopädie, Innerer Medizin und Neurologie. *Dtsch Arztebl.* (2006), 103, A–121–126
- [13] Hart D (2006) Vertrauen, Kooperation und Organisation. Probleme der Zusammenarbeit, der Übergabe und an Schnittstellen im Arzthaftungsrecht. In: Kern BR, Wadle E, Schroeder KP, Humaniora. *Medizin-Recht-Geschichte. Festschrift für Adolf Laufs zum 70. Geburtstag*, 843–879. Springer, Berlin
- [14] Hart D, (ed.) (2005) Ärztliche Leitlinien im Medizin- und Gesundheitsrecht. Recht und Empirie professioneller Normbildung. Nomos, Baden-Baden
- [15] Helou A et al., Methodische Standards der Entwicklung evidenz-basierter Leitlinien in Deutschland. Konsens zwischen Wissenschaft, Selbstverwaltung und Praxis. *Z Arztl Fortbild Qualitätssich.* (2000), 94, 330–339
- [16] BGH: Beschaffung eines Medikaments vor der Operation, zu Organisationspflichten dieser Art. *NJW.* (1991), 1543–1544
- [17] Hart D, „Organisationsaufklärung“ – Zum Verhältnis von Standardbehandlung, Organisationspflichten und ärztlicher Aufklärung. *MedR.* (1999), 17, 47–50
- [18] Hart D (2005) Leitlinien und Haftungsrecht – Inkorporation, Rezeption und Wissensbasis in Wissenschaft und Praxis. In: Hart D, *Klinische Leitlinien und Recht*, 81–103. Nomos, Baden-Baden
- [19] Dressler WD (2000) Ärztliche Leitlinien und Arzthaftung. In: Brandner HE, Hagen H, Stürner R, *Festschrift für Karlmann Geiß zum 65. Geburtstag.*, 379–384. Köln, Heymanns
- [20] Weidinger P, Leitlinien und ärztliche Berufshaftpflicht. *Internist.* (2001), 118–123
- [21] Diederichsen A (2005) Zur Bedeutung ärztlicher Leitlinien. In: Hart D, *Klinische Leitlinien und Recht*, 105–112. Nomos, Baden-Baden
- [22] Steffen E, Dressler WD (2002) Arzthaftungsrecht. Neue Entwicklungslinien in der BGH-Rechtsprechung. RWS Verlag, Köln
- [23] (2006) BGHZ 72, 132, 138 – Blinddarmoperation
- [24] (2006) BGHZ 88, 230 – Anfängeroperation
- [25] BGH: Beckenringfraktur. *NJW.* (2004), 2011 = *JZ* (2004), 1029 = *VersR* (2004), 909
- [26] Spickhoff A, Grober Behandlungsfehler und Beweislastumkehr. *NJW.* (2004), 2345
- [27] Katzenmeier C, Anmerkung BGH, 27.4.2004 – VI ZR 34/03 (S. 1029). Umkehr der Beweislast für die haftungsbegründende Kausalität. *JZ.* (2004), 20, 1029ff.
- [28] Katzenmeier C (2002) Arzthaftung. Mohr Siebeck, Heidelberg, 439ff.
- [29] BGH: grobe Verletzung der Pflicht zur Sicherungsaufklärung führt regelmäßig zur Umkehr der objektiven Beweislast für die haftungsbegründende Kausalität. *VersR.* (2005), 228

38b Evidenzbasierte Medizin und Recht: Sozialrecht

Felix Welti

38b.1 Einführung

Ob die Gesundheitsversorgung in Deutschland den Grundsätzen evidenzbasierter Medizin, Rehabilitation und Pflege folgt, hängt auch vom Sozialrecht ab. 90% der gesundheitlichen Leistungen werden im Rahmen des Sozialgesetzbuchs erbracht. Für das Anliegen evidenzbasierter Gesundheitsversorgung ist es daher wesentlich, ob die Akteure bei den Leistungsträgern und ihren Verbänden, bei den Leistungserbringern und schließlich auch die Leistungsberechtigten selbst die gesetzlichen Grundlagen als Auftrag verstehen, die gesundheitliche Versorgung nachweisgestützt zu gestalten oder ob sie Spannungen und Differenzen zwischen Recht und evidenzbasierter Gesundheitsversorgung sehen.

Dabei gibt es unterschiedliche Kritik: Während die einen den Auftrag der Sozialleistungsgesetze, nur notwendige Leistungen zu erbringen, für nicht hinreichend konkret halten und eine Konzentration der Leistungen auf das „wirklich Notwendige“ anmahnen, sehen andere in der Forderung nach evidenzbasierten Leistungen eine unberechtigte Schmälerung gesetzlicher Leistungsansprüche. Die Debatte ist verstärkt worden durch einen Beschluss des Bundesverfassungsgerichts vom 6. Dezember 2005, in dem dieses die bisherige Rechtsprechung des Bundessozialgerichts (BSG) als zu restriktiv kritisiert hat [1].

Im Folgenden wird dargestellt, wie die sozialrechtlichen Grundlagen der Gesundheitsversorgung in Deutschland sich auf die Forderung nach Evidenzbasierung auswir-

ken. Dabei konzentriert sich die Darstellung auf die gesetzliche Krankenversicherung. Rehabilitation und Pflege folgen vergleichbaren Grundsätzen, ohne jedoch über ein entsprechendes System von Institutionen zu verfügen [2; 3].

38b.2 Gesetzliche Grundlagen, insbesondere in der gesetzlichen Krankenversicherung (SGB V)

38b.2.1 Leistungsgrundsätze

Grundsatznorm für alle Leistungen der gesetzlichen Krankenversicherung ist § 2 SGB V:

§ 2 (1) Die Krankenkassen stellen den Versicherten die (...) Leistungen unter Beachtung des Wirtschaftlichkeitsgebots (§ 12) zur Verfügung (...). Behandlungsmethoden, Arznei- und Heilmittel der besonderen Therapierichtungen sind nicht ausgeschlossen. Qualität und Wirksamkeit der Leistungen haben dem allgemein anerkannten Stand der medizinischen Erkenntnisse zu entsprechen und den medizinischen Fortschritt zu berücksichtigen.

Mit dem Bezug auf den allgemein anerkannten Stand der medizinischen Erkenntnisse und den medizinischen Fortschritt werden hier der sozialrechtliche Leistungsanspruch und die Erkenntnisse der medizinischen Wissenschaft verknüpft. Der rechtlich

geforderte Standard entspricht im Sozialrecht wie im Haftungsrecht (vgl. Kap. 38a) dem medizinischen Standard [4–6].

Dass im Rahmen des medizinischen Standards grundsätzlich verschiedene medizinische und therapeutische Ansätze in Betracht kommen, wird daran deutlich, dass besondere Therapierichtungen nicht ausgeschlossen sind und der medizinische Fortschritt zu berücksichtigen ist. Beide Aussagen bedingen, dass sich der Anspruch nicht auf eine einzige Leistungsform konzentrieren muss, sondern auch minoritäre, besondere oder neue Behandlungsformen umfassen können. Damit wird jedoch nicht der Anspruch suspendiert, dass diese Leistungen dem Ziel der Krankenversicherung entsprechen müssen.

Die in der rechtswissenschaftlichen Literatur vertretene Ansicht, die gesonderte Erwähnung der besonderen Therapierichtungen befreie diese vom Erfordernis, einem allgemein anerkannten Stand der medizinischen Erkenntnisse zu entsprechen und lasse die reine Binnenanerkennung innerhalb der Therapierichtung genügen [7], findet in Wortlaut und Auslegung der Norm keine Stütze [8; 9]. Auch besondere Therapierichtungen müssen therapieren, sonst sind sie keine Therapierichtungen. Auch ihre Wirksamkeit muss überprüfbar sein. In einem richtigen Verständnis wendet sich die Norm gegen eine Beschränkung von Therapieformen auf eine bestimmte (vermutete) Wirkweise und (tradierte) Form. Sie steht insofern auch nicht, wie zum Teil behauptet wird, in einem Spannungsverhältnis zur Forderung, die Gesundheitsversorgung auf Nachweise (Evidenz) der Wirksamkeit und des Nutzens zu stützen (vgl. aber etwa [10]).

Neben dem medizinischen Standard wird für alle Leistungen das Gebot der Wirtschaftlichkeit festgeschrieben. Diese Ansprüche werden im Gesetz noch einmal explizit zusammengefasst:

§ 2 (4) Krankenkassen, Leistungserbringer und Versicherte haben darauf zu achten, dass die Leistungen wirksam und wirtschaftlich erbracht und nur im notwendigen Umfang in Anspruch genommen werden.

Mit dem Begriff der Wirksamkeit wird hier beschrieben, warum das Krankenversicherungsrecht den anerkannten Stand der medizinischen Erkenntnisse adaptiert. Wirksamkeit kann nur als generelle Wirksamkeit verstanden werden, da bei jeder Leistungsentscheidung eine Prognose ex ante erforderlich ist [11].

Das Gebot der Wirtschaftlichkeit wird am Beginn des Kapitels über die Leistungen im SGB V näher umschrieben:

§ 12 (1) Die Leistungen müssen ausreichend, zweckmäßig und wirtschaftlich sein; sie dürfen das Maß des Notwendigen nicht überschreiten. Leistungen, die nicht notwendig oder unwirtschaftlich sind, können Versicherte nicht beanspruchen, dürfen die Leistungserbringer nicht bewirken und die Krankenkassen nicht bewilligen.

Mit dem Gebot ausreichender Leistungen als Untergrenze und des Zweckmäßigen als Teil der Wirtschaftlichkeit wird im Gesetz deutlich gemacht, dass das Wirtschaftlichkeitsgebot keine Abstriche an der Wirksamkeit der Leistungen bedeuten soll [12]. Notwendig ist, was ausreichend und zweckmäßig ist, um den Leistungszweck zu erreichen. Dies soll so wirtschaftlich wie möglich geschehen, also mit dem geringstmöglichen Aufwand für das vorgegebene Ziel.

Die Grundsätze der Wirksamkeit, Notwendigkeit und Wirtschaftlichkeit gelten auch in allen anderen Bereichen sozialrechtlicher Gesundheitsleistungen. Die Notwendigkeit ist in den Rechtsgrundlagen der me-

dizinischen Rehabilitation (§ 26 I SGB IX) und der Pflegeversicherung (§ 29 I SGB XI) genannt. Die gesetzliche Unfallversicherung bezieht sich ausdrücklich auf Qualität und Wirksamkeit nach dem allgemein anerkannten Stand der medizinischen Erkenntnisse (§ 26 IV SGB VII). Wirtschaftlichkeit ist ein allgemeiner Grundsatz öffentlicher Leistungen (vgl. § 69 II SGB IV).

38b.2.2 Grundsätze der Leistungserbringung

Um die Leistungen der gesetzlichen Krankenversicherung zu konkretisieren, sind Verträge zwischen den Krankenkassen und den Leistungserbringern, insbesondere mit den Kassenärztlichen Vereinigungen als Vertretungen der Vertragsärzte und mit den Krankenhäusern vorgesehen. Die gesetzliche Grundsatznorm für dieses vertragliche Leistungserbringungsrecht wiederholt und präzisiert:

§ 70 (1) Die Krankenkassen und die Leistungserbringer haben eine bedarfsgerechte und gleichmäßige, dem allgemein anerkannten Stand der medizinischen Erkenntnisse entsprechende Versorgung der Versicherten zu gewährleisten. Die Versorgung der Versicherten muss ausreichend und zweckmäßig sein, darf das Maß des Notwendigen nicht überschreiten und muss in der fachlich gebotenen Qualität sowie wirtschaftlich erbracht werden.

Diese Norm ordnet die Verantwortlichkeit für die Versorgung der Versicherten insgesamt. Dabei geht es, richtig verstanden, um mehr als eine Summe von Einzelleistungen. Die Gewährleistung einer bedarfsgerechten und gleichmäßigen Versorgung umfasst auch Organisation und Koordination, wozu insbesondere gehört, fachliche Maßstäbe zu

ermöglichen und zu rezipieren. Die Pflicht zur bedarfsgerechten Versorgung setzt voraus, dass es Maßstäbe gibt, um einen Bedarf zu erkennen und zu definieren. Die Pflicht zur gleichmäßigen Versorgung setzt voraus, dass verglichen wird und werden kann. Diese Orientierung wird durch die gesonderte Bezugnahme der fachlich gebotenen Qualität verstärkt. Insgesamt wird deutlich, dass eine systematische wissenschaftliche Überprüfung der Wirksamkeit von Leistungen der Gesundheitsversorgung durch die gesetzlichen Vorgaben im Sozialrecht nicht nur zugelassen, sondern auch gefordert wird. Die Konkretisierung des sozialrechtlichen Anspruchsinhalts im einzelnen Fall ist regelmäßig Sache des Arztes, der als Vertragsarzt und im Krankenhaus die Leistungsentscheidung nach medizinischem Standard trifft.

38b.2.3 Der Gemeinsame Bundesausschuss

Um auch das Wirtschaftlichkeitsgebot zur Geltung zu bringen und eine bedarfsgerechte, gleichmäßige Leistung in der fachlich gebotenen Qualität sicherzustellen, ist ein Rahmen für die ärztliche Entscheidung erforderlich, mit der dieser die Leistungspflicht der Krankenkasse konkretisiert. Der Gesetzgeber hat den Gemeinsamen Bundesausschuss (G-BA) beauftragt, Richtlinien über die Gewähr für eine ausreichende, zweckmäßige und wirtschaftliche Versorgung der Versicherten zu beschließen und ihn ermächtigt, die Erbringung und Verordnung von Leistungen oder Maßnahmen verbindlich einzuschränken oder auszuschließen, wenn nach dem allgemein anerkannten Stand der medizinischen Erkenntnisse der diagnostische oder therapeutische Nutzen, die medizinische Notwendigkeit oder die Wirtschaftlichkeit nicht nachgewiesen sind (§§ 92 I, 91 IX SGB V) (s. Kap. 36).

Eine besondere Regelung haben neue Untersuchungs- und Behandlungsmethoden

erfahren. Während Methoden im Krankenhaus angewandt werden können, solange der G-BA nicht in einer Richtlinie negativ festgestellt hat, dass sie für die ausreichende, zweckmäßige und wirtschaftliche Versorgung der Versicherten nicht erforderlich sind (§ 137c SGB V), sind neue Untersuchungs- und Behandlungsmethoden in der vertragsärztlichen Versorgung nur zulässig, wenn der G-BA in Richtlinien dazu positive Empfehlungen abgegeben hat (§ 135 Abs. 1 SGB V). **Während also für neue Methoden im Krankenhaus eine Erlaubnis mit Verbotsvorbehalt besteht, gilt in der ambulanten vertragsärztlichen Versorgung ein Verbot mit Erlaubnisvorbehalt.** Damit wird in restriktiver Weise der medizinische Fortschritt berücksichtigt, seine ungeprüfte Anwendung insbesondere im vertragsärztlichen Bereich aber als Gefahr für die Leistungsgrundsätze einer wirksamen und wirtschaftlichen, bedarfsgerechten und gleichmäßigen Versorgung eingestuft.

Noch präziser sind die gesetzlichen Vorgaben für die versorgungspolitische Innovation der strukturierten Behandlungsprogramme bei chronischen Krankheiten (§ 137f SGB V). Der G-BA soll Krankheiten empfehlen, die geeignet sind, um solche strukturierten Behandlungsprogramme der Krankenkassen zu entwickeln. Ein Kriterium für den G-BA bei der Auswahl der geeigneten chronischen Krankheiten soll die Verfügbarkeit von evidenzbasierten Leitlinien sein.

Eine gewisse Verschiebung dieses Verhältnisses von Gestaltung und Rezeption hat der Gesetzgeber durch die Gründung des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG, § 139a SGB V) angelegt. Dieses Institut wird vom G-BA getragen und soll fachlich unabhängig und wissenschaftlich arbeiten. Zu seinen Aufträgen gehören die Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissensstandes zu diagnostischen und therapeutischen Verfahren bei ausgewählten

Krankheiten sowie Bewertungen evidenzbasierter Leitlinien für die epidemiologisch wichtigsten Krankheiten. Der Aufgabenbeschreibung liegt die Annahme zugrunde, dass entsprechende Erkenntnisse über die Gesundheitsversorgung nicht hinreichend vorliegen, also nicht nur rezipiert, sondern zunächst produziert werden müssen. Mit dem gesetzlichen Auftrag an den G-BA wird zugleich anerkannt, dass diese Aufgabe nur fachlich unabhängig und wissenschaftlich erfüllt werden kann. Dem G-BA wird gesetzlich aufgegeben, die Empfehlungen des IQWiG im Rahmen seiner Aufgabenstellung zu berücksichtigen (§ 139a IV 2 SGB V).

Im G-BA sind die Kassenärztlichen Bundesvereinigungen, die Deutsche Krankenhausesellschaft und die Bundesverbände der Krankenkassen und Ersatzkassen stimmberechtigt. Beratend nehmen Vertreter der für die Wahrnehmung der Patientinnen- und Patientenrechte und der Selbsthilfe chronisch kranker und behinderter Menschen maßgeblichen Organisationen teil (§ 140f SGB V). Politisch und verfassungsrechtlich ist stark umstritten, ob der Gesetzgeber den G-BA mit dieser für die Gesundheitsversorgung im Rahmen der Krankenversicherung bedeutsamen Aufgabe betrauen darf und sollte.

Der politische und rechtliche Streit um die Kompetenzen des G-BA muss alle interessieren, die nach der Verankerung von Evidenzbasierung im deutschen Gesundheitswesen fragen, denn er betrifft die Frage, wer zu verbindlichen Entscheidungen über die Gesundheitsversorgung berufen sein soll. Dabei ist die Frage, wer entscheidet, mit der Frage verknüpft, nach welchen Maßstäben und Methoden entschieden werden soll.

Betrachtet man den gesetzlichen Auftrag des G-BA, seine selbstgesetzten Verfahrensregeln und die gesetzliche Verknüpfung mit der Tätigkeit des IQWiG, so wird man feststellen, dass sich der G-BA vor allem sachlich als das Gremium legitimieren muss, das in

der Lage ist, den anerkannten Stand der medizinischen Erkenntnisse zu erkennen und ihn mit den Anforderungen des Wirtschaftlichkeitsgebots so zu verknüpfen, dass die notwendigen Leistungen der Krankenbehandlung konkretisiert werden. Dabei kann nicht der einzelne Behandlungsfall geregelt, sondern nur ein Korridor fachlich und gesetzlich vertretbarer therapeutischer Leistungen beschrieben werden.

Wäre der G-BA nicht fachlich legitimiert oder würde er seine fachliche Aufgabe nicht hinreichend erfüllen, könnte er sich kaum alleine durch personale demokratische Legitimation der sozialen Selbstverwaltung für seine Aufgabe qualifizieren. Demokratisch legitimiert ist primär der Gesetzgeber. Dessen Auftrag, wirksame und wirtschaftliche Leistungen der Krankenbehandlung gleichmäßig und bedarfsgerecht bereitzustellen, soll der Bundesausschuss kraft fachlicher Legitimation ausfüllen. Hierzu bedarf er der Methoden evidenzbasierter Gesundheitsversorgung.

Es ist demokratisch und rechtsstaatlich, wenn der Gesetzgeber entscheidet, welche gesundheitlichen Leistungen solidarisch finanziert und beansprucht werden können. Dies entspricht dem Vorrang und Vorbehalt des Gesetzes im demokratischen Rechtsstaat. Doch auch der rechtsstaatliche Gesetzesvorbehalt setzt voraus, dass eine angemessene Arbeitsteilung zwischen Gesetzgebung und Gesetzesanwendung gefunden wird und fachliche Entscheidungen fachlich unabhängig getroffen werden können. Je stärker aber auch wertende Elemente in fachlichen Entscheidungen enthalten sind, desto drängender wird die Frage nach der Legitimation derer gestellt, die sie treffen. Entscheidungen über die gesundheitliche Versorgung nach einer fachlich erarbeiteten, wissenschaftlich tragfähigen und allgemein verständlichen Methode auf Nachweise der Wirksamkeit (Evidenz) zu stützen, ist ein wichtiger Beitrag dazu, das Legitimationsproblem zu lösen. In-

sofern ist ein weiterer und intensiver Dialog zwischen Rechtswissenschaft, Medizin und demokratischer Öffentlichkeit über das Verfahren der Entscheidung im Gemeinsamen Bundesausschuss einzufordern.

38b.3 Verfahrensordnung des G-BA

Das Verfahren des G-BA ist in seiner Verfahrensordnung geregelt, die am 1. Oktober 2005 in Kraft getreten ist. Darin ist ein Bewertungsverfahren enthalten, das den allgemein anerkannten Stand der medizinischen Erkenntnisse zu Nutzen, Notwendigkeit und Wirtschaftlichkeit einer Methode feststellen soll. Die Bewertung wird öffentlich angekündigt und Sachverständigen der medizinischen Wissenschaft und Praxis und Fachverbänden wird Gelegenheit zur Stellungnahme gegeben (§ 13 VerfO). Die vorliegenden Unterlagen werden nach Evidenzstufen klassifiziert (§ 18 VerfO). Ein Unterausschuss bewertet die Qualität und Aussagekraft der Unterlagen nach der Studienqualität und ihrer Übertragbarkeit auf die Versorgungsrealität (§ 18 Abs. 5 und 6 VerfO). Dabei sollen insbesondere auch die Verwendung patientenrelevanter Zielgrößen (wie z.B. Mortalität, Morbidität, Lebensqualität), Versorgungsaspekte von Alter, biologischem und sozialem Geschlecht sowie lebenslagenspezifische Besonderheiten, besondere Belange behinderter und chronisch kranker Menschen und die eingesetzten Maßnahmen zur Vermeidung von verzerrten Studienergebnissen berücksichtigt werden (§ 18 Abs. 7 VerfO).

Für die Gesamtbewertung im Gemeinsamen Bundesausschuss gilt § 20 VerfO. Diese Norm ist wegen der Anforderungen an die Evidenzstufen bei der Erarbeitung der Verfahrensordnung zwischen Bundesausschuss und Bundesministerium für Gesundheit umstritten gewesen. Auch die Arbeitsgemeinschaft Wissenschaftlich Medizinischer Fachgesellschaften (AWMF) hatte sich kritisch zu

einer Überbewertung der Evidenzstufe 1 geäußert. Die Kritik ist berücksichtigt worden. Die Norm lautet nun:

§ 20 (1) Das Beschlussgremium bewertet in einem umfassenden Abwägungsprozess unter Einbeziehung der wissenschaftlichen Erkenntnisse, insbesondere der nach Evidenzkriterien ausgewerteten Unterlagen.

(2) Der Nutzen einer Methode ist durch qualitativ angemessene Unterlagen zu belegen. Dies sollten, soweit möglich, Unterlagen der Evidenzstufe 1 mit patientenbezogenen Endpunkten (z.B. Mortalität, Morbidität, Lebensqualität) sein. Bei seltenen Erkrankungen, bei Methoden ohne vorhandene Alternative oder aus anderen Gründen kann es unmöglich oder unangemessen sein, Studien dieser Evidenzstufe durchzuführen oder zu fordern. Soweit qualitativ angemessene Unterlagen dieser Aussagekraft nicht vorliegen, erfolgt die Nutzen-Schaden-Abwägung einer Methode auf Grundlage von Unterlagen einer niedrigeren Evidenzstufe. Die Anerkennung des medizinischen Nutzens einer Methode auf Grundlage von Unterlagen einer niedrigeren Evidenzstufe bedarf jedoch – auch unter Berücksichtigung der jeweiligen medizinischen Notwendigkeit – zum Schutz der Patienten umso mehr einer Begründung, je weiter von der Evidenzstufe 1 abgewichen wird. Dafür ist der potenzielle Nutzen einer Methode insbesondere gegen die Risiken der Anwendung beim Patienten abzuwägen, die mit einem Wirksamkeitsnachweis geringerer Aussagekraft einhergehen.

(3) Die Bewertung der medizinischen Notwendigkeit erfolgt im Versorgungskontext unter Berücksichtigung der Relevanz der medizinischen Problematik, Verlauf und Behandelbarkeit der Erkran-

kung und insbesondere der bereits in der GKV-Versorgung etablierten diagnostischen und therapeutischen Alternativen. Maßstab ist dabei auch die von der Anwendung der Methode bereits erzielte oder erhoffte Verbesserung der Versorgung durch die Gesetzliche Krankenversicherung unter Berücksichtigung der mit der Erkrankung verbundenen Einschränkung der Lebensqualität und den besonderen Anforderungen an die Versorgung spezifischer Patientengruppen unter Berücksichtigung der Versorgungsaspekte von Alter, biologischem und sozialem Geschlecht sowie der lebenslangenspezifischen Besonderheiten.

Die mehrfache Betonung von Abwägung, Gewichtung und Berücksichtigung auch sozialer Faktoren zeigt, dass die Gesamtbewertung durch den G-BA zwar auf der Grundlage evidenzbasierter Studien nach gesetzlichen und selbstgesetzten Kriterien erfolgt, aber eigene Spielräume zu gesundheits- und sozialpolitischer Wertung bleiben. Die Orientierung des G-BA an den Kriterien Evidenzbasierter Medizin verdeutlicht zwar dessen fachliche Legitimation, vermag jedoch die verfassungsrechtliche und politische Frage nicht abschließend aufzulösen, wie viel Spielraum seinen Akteuren bei der Bewertung der so gewonnenen Erkenntnisse verbleiben darf, kann und soll. Festzuhalten ist aber, dass bei korrekter Einhaltung der Verfahrensordnung die Furcht in Politik und Öffentlichkeit vermindert werden kann, der G-BA treffe versorgungspolitische Entscheidungen nach eigener Willkür, nach dem Prinzip größtmöglicher Einsparung oder zur Fernhaltung unliebsamer Konkurrenz im Gesundheitswesen.

38b.4 Rechtsprechung der Sozialgerichtsbarkeit

Es gibt in Rechtswissenschaft und Rechtsprechung kaum Zweifel daran, dass Nachweise zu fordern sind, ob eine nach Krankenversicherungsrecht zu beanspruchende Leistung den gesetzlichen Kriterien der Wirksamkeit, Wirtschaftlichkeit und Notwendigkeit entspricht. Rechtliche Grundregel ist, dass, wer einen Anspruch erhebt, beweisen muss, dass die Voraussetzungen dieses Anspruchs vorliegen. Beweiserleichterungen werden akzeptiert, wenn Nachweise aufgrund von Umständen erschwert sind, die der Anspruchsteller nicht zu vertreten hat und insbesondere dann, wenn diese Umstände in der Sphäre des Antragsgegners liegen. Diese rechtliche Grundstruktur ist auch jedem Streit über die Berücksichtigung von Evidenz in der Rechtswissenschaft zugrunde zu legen.

Neben der Strenge von Evidenzanforderungen ist umstritten, wer vorgelegte Evidenz bewerten kann. Das aus vielen gerichtlichen Verfahren bekannte Problem der Abgrenzung zwischen fachlicher Expertise des Sachverständigen und rechtlicher Würdigung und Bewertung durch das Gericht stellt sich auch hier. Juristinnen und Juristen wollen vermeiden, Streitigkeiten zu entscheiden, die in der Substanz nicht rechtlich sind, sondern nur den Regeln einer anderen Disziplin folgen. Eine schärfere Abgrenzung wird dann möglich, wenn es gelingt, formale Anforderungen an fachliche Entscheidungsprozesse verbindlich so zu fixieren, dass ein rechtlich kontrollierbarer Bereich entsteht, von dem der genuin fachliche, mit rechtlichen Mitteln nicht überprüfbare Bereich klar geschieden ist.

Das Bundessozialgericht (BSG) hat zur Bestimmung des anerkannten Stands der medizinischen Erkenntnisse in mehreren Entscheidungen der letzten Jahre die folgende Formel benutzt, in der es der Sache nach auf Kriterien der evidenzbasierten Gesundheitsversorgung zurückgreift [13–15]:

„Die Behandlung muss sich in einer für die sichere Beurteilung ausreichenden Zahl von Fällen als erfolgreich erwiesen haben, was in der Regel durch wissenschaftlich einwandfrei geführte Statistiken belegt sein muss. (...) Da es auf den Nachweis genereller Wirksamkeit ankommt, kann die Leistungspflicht der Krankenkasse nicht damit begründet werden, dass die Therapie im konkreten Einzelfall erfolgreich gewesen sei und es unter der Behandlung zu einer Besserung des Gesundheitszustandes gekommen sei.“

Weiterhin hat das BSG in den letzten Jahren hohe Anforderungen daran gestellt, die Richtigkeit von Entscheidungen oder Nicht-Entscheidungen des Bundesausschusses in Frage zu stellen. Hierfür akzeptierte das BSG bisher nur Verfahrensfehler beim G-BA (Systemversagen). Ein solches Systemversagen kann auch dann vorliegen, wenn eine neue Methode bereits eine solche fachliche Resonanz gefunden hat, dass der G-BA sich mit ihr hätte befassen müssen.

Die juristische Kritik an der Rechtsprechung des BSG sah insbesondere rechtsstaatliche Grundsätze dadurch gefährdet, dass der gerichtlich nicht kontrollierte Spielraum des G-BA zu weit ausgedehnt worden sei und damit die Grundrechte Betroffener verletzt würden.

38b.5 Die Entscheidung des Bundesverfassungsgerichts vom 6.12.2005

In einem Beschluss vom 6.12.2005 [1] hat das BVerfG eine Entscheidung des BSG zur Konkretisierung des allgemein anerkannten Stands der medizinischen Erkenntnisse aufgehoben. Im Streit stand die Leistungspflicht für die Behandlung der Duchenne'schen

Muskeldystrophie durch hochfrequente Schwingungen (Bioresonanztherapie). Der Beschwerdeführer hatte eine Besserung seines Zustands geltend gemacht, die nach Auffassung der Krankenkasse nicht auf die Therapie zurückzuführen war. Das BSG hatte eine Leistungspflicht schon deshalb abgelehnt, weil der G-BA keine positive Entscheidung über die Bioresonanztherapie getroffen hatte.

Das Bundesverfassungsgericht (BVerfG) hat die Entscheidung des BSG als Verstoß gegen die Grundrechte des Patienten aufgehoben und den Fall zur erneuten Entscheidung an das BSG verwiesen, wo er wegen langer Verfahrensdauer durch Vergleich beendet wurde. Das BVerfG führte aus, dass es nicht mit der Handlungsfreiheit des Versicherten in Verbindung mit dem Sozialstaatsgebot und mit der Schutzpflicht des Staates für Leben und Gesundheit vereinbar sei, ihm bei einer lebensbedrohlichen oder regelmäßig tödlichen Erkrankung eine Behandlung vorzuenthalten.

Die Kernsätze über die Beweisanforderungen, welche die sozialgerichtliche Rechtsprechung wird umsetzen müssen, lauten [16]:

„Dabei muss allerdings die vom Versicherten gewählte Behandlungsmethode eine auf Indizien gestützte, nicht ganz fern liegende Aussicht auf Heilung oder wenigstens auf eine spürbare Einwirkung auf den Krankheitsverlauf versprechen. (...) Die im Streitfall vom Versicherten angerufenen Sozialgerichte haben in solchen Fällen, gegebenenfalls mit sachverständiger Hilfe, zu prüfen, ob es für die vom Arzt nach gewissenhafter fachlicher Einschätzung vorgenommene oder von ihm beabsichtigte Behandlung ernsthafte Hinweise auf einen nicht ganz entfernt liegenden Erfolg der Heilung oder auch nur auf eine spürbare positive Einwirkung auf den Krankheitsverlauf im konkreten Einzelfall gibt (...). Solche Hinweise auf einen individuellen

Wirkungszusammenhang können sich aus dem Gesundheitszustand des Versicherten im Vergleich anderer, in gleicher Weise erkrankten, aber nicht mit der in Frage stehenden Methode behandelten Personen ergeben sowie auch mit dem solcher Personen, die bereits auf diese Weise behandelt wurden oder behandelt werden. (...) Weitere Bedeutung kommt der fachlichen Einschätzung der Wirksamkeit der Methode im konkreten Einzelfall durch die Ärzte des Erkrankten zu, die die Symptome seiner Krankheit behandeln.“

Mit der Entscheidung des BVerfG ist der G-BA wieder stärker an die rechtliche Handlungslogik rückgebunden worden. Zu dieser gehören die gerichtliche Überprüfbarkeit hoheitlicher Entscheidungen und das Bemühen um Einzelfallgerechtigkeit. Eine konsequente Orientierung an überindividuellen Evidenzen möglichst hoher Stufen ist damit in Frage gestellt. Die Entscheidung des BVerfG kann aber nicht als Absage an die Berücksichtigung wissenschaftlicher Evidenz verstanden werden, wie sich aus der klaren Bezugnahme auf Vergleichsstudien ergibt. Möglicherweise ist sie ein Impuls, eine Leistungspflicht der gesetzlichen Krankenversicherung für Heilversuche in bestimmtem Umfang zu normieren [17].

Das BSG hat in ersten Entscheidungen festgehalten, dass auch nach den Maßstäben des BVerfG von einer Methode objektivierbare wissenschaftliche Erkenntnisse zu verlangen sind, nach denen bei Berücksichtigung des gebotenen Wahrscheinlichkeitsmaßstabes sowohl eine abstrakte als auch eine konkret-individuelle Abwägung von Chancen und Risiken ergeben, dass die Chancen die Risiken überwiegen. Zudem müssten sich neue Methoden einem Vergleich mit anerkannten Methoden hinsichtlich Eignung, Erforderlichkeit und Wirtschaftlichkeit aus-

setzen. Je schwerwiegender die Erkrankung ist, desto geringere Anforderungen sind an die ernsthaften Hinweise auf einen nicht ganz fern liegenden Behandlungserfolg zu stellen [18].

38b.6 Schluss

Welche Folgerungen aus der Entscheidung des BVerfG für die sozialgerichtliche Rechtsprechung und die Arbeit des G-BA gezogen werden, wird in nächster Zeit deutlich werden. Dabei ist es Aufgabe der an evidenzbasierter Gesundheitsversorgung interessierten Wissenschaft, darzulegen, welchen Nutzen ihre Methoden und Erkenntnisse für eine für insgesamt zweckmäßige Verteilung von Mitteln im Gesundheitswesen und für gerechte Entscheidungen im Einzelfall zu liefern vermögen.

In der rechtswissenschaftlichen Literatur und selbst in der Entscheidung des Bundesverfassungsgerichts erscheint dabei immer wieder die „Schulmedizin“ als Begriff. Obwohl im Krankenversicherungsrecht nirgendwo in Bezug genommen, ist die „Schulmedizin“ im Diskurs immer wieder die Chiffre für eine im negativen Sinne dogmatische, an überkommenen Auffassungen und Methoden orientierte Medizin. Die Vertreter einer stärker Evidenzbasierten Medizin müssen deutlich machen, dass ihr methodischer Ansatz sich gerade kritisch gegen ungeprüfte Überlieferungen wendet und dass sie einen wichtigen Beitrag zu der von vielen Menschen gewünschten Neuausrichtung des Gesundheitswesens an Lebensqualität, Teilhabe und Effizienz leisten können.

Literatur

- [1] BVerfG vom 6.12.2005, Az. 1 BvR 347/98. NJW. (2006), 891
- [2] Igl G, Evidenzbasierte Rehabilitation und Pflege im Spiegel des Sozialrechts. Z Arztl Fortbild Qualitätssich. (2005), 99, 25–33

- [3] Welti F, Raspe H, Der rechtliche Rahmen der evidenzbasierten Rehabilitation. Dtsch Rentenvers. (2005), 576
- [4] Neumann V, Nicklas-Faust J, Evidenzbasierte Leitlinien als Maßstab ärztlichen Handelns in medizinischer und rechtlicher Sicht. Recht soz Dienst Einrichtungen. (2006), 60, 23–49
- [5] Hart D (ed.) (2005) Ärztliche Leitlinien im Medizin- und Gesundheitsrecht. Recht und Empirie professioneller Normbildung. Nomos, Baden-Baden
- [6] Mengel C (2004) Sozialrechtliche Rezeption ärztlicher Leitlinien. Nomos, Baden-Baden
- [7] Zuck R, Was wird aus der Akupunktur. NJW. (2001), 869
- [8] Welti F, Raspe H, Evidenz und Akzeptanz von Medizin und Recht. NJW. (2002), 850–874
- [9] Welti F, Der sozialrechtliche Rahmen ärztlicher Therapiefreiheit. Gesundheitsrecht. (2005), 1–12
- [10] Busse A, Eine Dogmatik der besonderen Therapierichtungen des SGB V. Sozialgerichtsbarkeit. (2000), 61–68
- [11] Noftz W (2000) Kommentierung zu § 12 SGB V. In: SGB V, 46. Ergänzungs-Lieferung I/2000, Rz 27
- [12] Noftz W (2000) Kommentierung zu § 12 SGB V. In: SGB V, 46. Ergänzungs-Lieferung I/2000, Rz 23
- [13] BSG vom 5.7.1996-1 RK 6/95 (Remedacien). BSGE. (1996), 76, 194–199
- [14] BSG vom 16.6.1999-B 1 KR 4/98 R (Kozijavkin). Sozialgerichtsbarkeit. (2000), 322–325
- [15] BSG vom 18.5.2005-B 1 KR 21/02 R (Immucothel). N Z Sozialrecht. (2005), 308
- [16] BVerfG vom 6.12.2005, Az. 1 BvR 347/98. Randziffer 64, 66. www.bverfg.de (14.12.06)
- [17] Francke R, Hart D, Die Leistungspflicht der gesetzlichen Krankenversicherung für Heilversuche-Zugleich eine Besprechung der Entscheidung des BVerfG vom 6.12.2005 und die Skizzierung eines Regulierungsmodells. Medizinrecht. (2006), 131–138
- [18] BSG vom 4.4.2006 – B1 KR 7/05 R, SozR 4-2500 § 31 Nr. 4; BSG vom 7.11.2006 – B 1 KR 24/06 R

V Aus-, Weiter- und Fortbildung

39	Ausbildung	413
	<i>Martin Bergold, Tobias Weberschock</i>	
40	Weiter- und Fortbildung	421
	<i>Norbert Donner-Banzhoff, Olaf Weingart, Gerd Burmester, Susanne Weinbrenner, Hans Wille, Regina Kunz</i>	
41	Schulung von medizinischen Laien	427
	<i>Bettina Berger, Nicole Skoetz, Olaf Weingart</i>	

39 Ausbildung

Martin Bergold*, Tobias Weberschock* (*equal contributor)

Szenario

Im Internet steht, dass einmalig Azithromycin bei einer frühen Syphilis ausreichend ist¹. Stimmt das denn?“

Auf diese Art und Weise werden Sie eines Mittags von einem Frankfurter Medizinstudierenden gefragt. Anstatt einfach eine Antwort zu geben, entschließen Sie sich für etwas Lehre in Evidenzbasierter Medizin (EbM).

Zielgruppe

Alle, die Ausbildung in EbM betreiben möchten oder bereits betreiben.

Lernziele

- ▲ Sie können unterscheiden zwischen evidenzbasierter Lehre und Lehre der EbM.
- ▲ Sie können 2 verschiedene Szenarien der EbM-Ausbildung beschreiben.
- ▲ Sie können die Besonderheiten der schrittweisen Ausbildung in EbM erklären.
- ▲ Sie können 12 Tipps für einen gelungeneren Unterricht aufzählen.
- ▲ Sie können die Eigenschaften von formativen und summativen Prüfungen definieren.

39.1 Sie können unterscheiden zwischen evidenzbasierter Lehre und Lehre der EbM

Das oben beschriebene Szenario kann grundsätzlich durch zwei verschiedene Möglichkeiten gelöst werden.

Die erste Möglichkeit wäre die direkte Antwort auf die gestellte Frage. Dies ist die

einfachste und unkomplizierteste Art, Lehre zu betreiben. Ohne großen Aufwand und innerhalb kürzester Zeit würde der Studierende eine neue medizinische Information erhalten. Dabei wird selbstverständlich vorausgesetzt, dass der Inhalt der Antwort auch evidenzbasiert ist. Solches Vermitteln von aktueller, wissenschaftlicher und evidenzbasierter Information in der medizinischen Ausbildung möchten wir als „**evidenzbasierte Lehre**“ bezeichnen.

Die zweite Möglichkeit wäre die „**Lehre in Evidenzbasierter Medizin**“. Dabei sollen dem Studierenden das nötige Hintergrundwissen und die Fertigkeiten der Evidenzbasierten Medizin vermittelt werden. Zugegebenermaßen würde bei dieser Variante der Studierende seine Antwort nicht unmittelbar erhalten – je nach Szenario müsste er sich die Antwort sogar selbst erarbeiten. Auf längere Sicht jedoch würde jeder Studierende ungenutzt von diesem Weg profitieren.

Die 9. Novelle der ärztlichen Approbationsordnung (ÄAppO) fordert als grundlegende Ausbildungsziele:

- ▲ eine fundierte wissenschaftliche und praktische Ausbildung,
- ▲ die Befähigung zur eigenverantwortlichen und selbstständigen Berufsausübung und
- ▲ die Befähigung zur eigenen ständigen Fortbildung [1].

Die evidenzbasierte Lehre allein erfüllt nur die erste Forderung. Die Befähigung zur ei-

1 MMA 2005, Ärztemagazin 40/2005

genverantwortlichen/selbstständigen Berufsausübung und zur selbstständigen Fortbildungsfähigkeit sind jedoch explizite Ausbildungsinhalte der EbM.

Um also alle drei Forderungen zu erfüllen, muss auch die Lehre in Evidenzbasierter Medizin betrieben werden.

39.2 Sie können 2 verschiedene Szenarien zur schrittweisen Ausbildung beschreiben

Zwei Unterrichtsformen eignen sich besonders für die Ausbildung in Evidenzbasierter Medizin. Zum einen ist dies das Bedside-Teaching, zum anderen das Seminar.

Beim Bedside-Teaching werden dem Studierenden echte klinische Szenarien dargeboten (Station, echte Patienten etc.). Basierend auf einer realen klinischen (beantwortbaren) Frage haben die Studierenden im Sinne einer „Educational Prescription“ die Aufgabe, alle Schritte der EbM „am Patientenbett“ zu durchlaufen [2].

Ein großer Vorteil dieses Modells ist der äußerst nahe Bezug zur medizinischen Praxis. Auch die Möglichkeit, gewonnene evidenzbasierte Erkenntnisse direkt auf den individuellen Patienten übertragen zu können, ist von großem Wert.

Solch ein Modell kommt daher auch bei den Studierenden gut an. Epidemiologische und biometrische Inhalte, die meist von Studierenden als langweilig und trocken empfunden werden, bekommen auf einmal einen praktischen und realen Bezug. Plötzlich werden epidemiologische Zusammenhänge erkannt und deren Bedeutung anders gewertet.

Auch das Seminar wird von den Studierenden durchaus positiv aufgenommen. In kleineren Gruppen wechseln sich hierbei Eigenarbeiten, Gruppenarbeiten und Vortragsteile ab. Statt mit realen klinischen Szenarien wird der Studierende mit fiktiven Szenarien

konfrontiert. Dabei sollen – wie im Bedside-Teaching – ebenfalls die Schritte der EbM durchlaufen werden.

In diesem Modell spielen fiktive Fallbeispiele, auch Paper-Cases genannt, eine große Rolle. Diese schaffen auch den zur Motivation notwendigen Bezug, sind dabei allerdings kalkulierbarer, da sie im Vorhinein erstellt und ausprobiert werden können.

Für welche Unterrichtsform Sie sich auch immer für die Ausbildung in EbM entscheiden – die Studierenden sollten so oft wie möglich die Gelegenheit erhalten, Erlerntes auch in die Praxis umsetzen zu können. Denn die Erkenntnis, dass der Inhalt der EbM für das Arztsein unabdingbar ist, stellt sich nur ein, wenn die Studierenden den Bezug der EbM zur praktischen Medizin selbst verspüren.

39.3 Sie können die Besonderheiten der schrittweisen Ausbildung in EbM erklären

Kommen wir nun zu unserem Fallbeispiel zurück. Ihnen fällt sicherlich auf, dass die medizinische Fragestellung des Studierenden etwas zu ungenau ist. Gibt es ein Vergleichsmittel? Was ist mit „ausreichend“ gemeint?

Unsere Erfahrung hat gezeigt, dass ein Großteil der Studierenden oft nicht in der Lage ist, eine präzise Fragestellung bei einem vorhandenen medizinischen Problem zu erstellen. So banal dies auch klingt: Die Bildung einer präzisen medizinischen Fragestellung mit einer genau definierten Zielsetzung steht am Anfang der Ausbildung in EbM.

Zeigen Sie dem Studierenden, dass die Bildung einer simplen (aber präzisen) vierteiligen Frage das weitere wissenschaftliche Vorgehen erleichtern kann. Geben Sie dem Studierenden ein paar Fallbeispiele in die Hand, zu denen er/sie vierteilige Fragen erstellen soll – so lernt er rasch, wie man eine

präzisere Frage stellt. Derselbe Studierende aus dem oben genannten Szenario wird Ihnen nun die folgende Frage stellen: „Ist bei der frühen Syphilis die einmalige orale Gabe Azithromycin oder die standardmäßige einmalige i.m.-Injektion Benzathinpenicillin geeigneter, um eine serologisch kontrollierte Abheilung zu erreichen?“

Jetzt geht es darum, dem Studierenden zu zeigen, wie man im Internet nach geeigneter Information sucht (s. Kap. 20). PubMed eignet sich dabei hervorragend, Studierenden die medizinische Internetrecherche beizubringen. Sicherlich sind andere Suchoberflächen wie die Cochrane Library (s. Kap. 43) in diesem Falle auch sehr geeignet, aber warum dem Studierenden nicht gleich die Handhabung der meistgenutzten und gleichzeitig frei verfügbaren medizinischen Suchoberfläche zeigen?

Lässt man Studierende nun ohne weitere Anleitung suchen, so werden Sie merken, dass diese schnell an ihre Grenzen gelangen. Nicht wenige Studierende tendieren dazu, die rund 23 000 Treffer, die sie bei der Eingabe des Suchwortes „**syphilis**“ erhalten, manuell zu durchsuchen. Dass man unter diesen Voraussetzungen schnell den nötigen Antrieb verliert, dürfte offensichtlich sein.

Bringen Sie also dem Studierenden zuerst die Anwendung von gängigen Boole'schen Operatoren wie „AND“ oder „OR“ bei und zeigen Sie ihm den Unterschied zwischen einer spezifischen und einer sensitiven Suche. So können Sie vermeiden, dass der Studierende später entweder zu viele oder zu wenige Suchwörter eingibt und somit die Recherche misslingt.

Zwar ist der Studierende nun über die Tatsache erfreut, dass er durch die Eingabe des Suchterms „**syphilis AND azithromycin AND penicillin**“ nur 31 Treffer erhalten hat. Ihm wird jedoch schnell klar, dass auch das Lesen von 31 Titelüberschriften und/oder Abstracts zu einer Last werden kann. Daher müssen andere Suchfunktionen wie „**limits**“

oder „**tags**“ und deren adäquate Anwendung ebenfalls vermittelt werden. Eine MeSH-Suche ist auch ein essenzieller Bestandteil der Internetrecherche, jedoch ist er eher für Studierende geeignet, die in der Suche schon routinierter sind.

Der Suchterm „**syphilis AND azithromycin AND penicillin, limits: Randomized Controlled Trial**“ liefert nun 5 Treffer. Nach gründlicher Erwägung entscheidet sich der Studierende für die genauere Betrachtung der Studie an erster Stelle [3]. Denken Sie daran, den Studierenden auch den Weg zu demonstrieren, wie man von dem Abstract bei PubMed zu dem Volltext im PDF-Format gelangt (z.B. durch die elektronische Zeitschriftenbibliothek der eigenen Universität).

Nun steht der Studierende vor einer erneuten Schwierigkeit: Wie soll er die Studie, die zudem noch in Englisch verfasst ist, lesen, verstehen und obendrein auch noch bewerten? Oftmals ist es vielmehr die Angst vor der englischen Sprache als der Fachinhalt, der die Studierenden vor der Bearbeitung von Studien zurückschrecken lässt.

Ihre Aufgabe ist es, ihm die Angst vor fremdsprachigen Texten zu nehmen. Geben Sie ihm eine Art Anleitung, wie er die Studie lesen soll. Machen Sie ihm klar, dass es nur notwendig ist, die relevanten Teile zu verstehen.

„Nebenbei“ können (und sollen) Sie ihm die Validitätskriterien von Therapiestudien nahe bringen (s. Kap. 12–14). Die Inhalte der Randomisation, der Verblindung, der verschiedenen systematischen Fehler („bias“) und die anderen Punkte werden im Allgemeinen schnell verstanden. Vereinfachen können Sie sich die Vermittlung der einzelnen epidemiologischen Grundbegriffe, wenn Sie zur Erklärung ihrer Bedeutung plastische Beispiele verwenden.

Lediglich die Vermittlung des „Intention-to-treat“ (ITT) kann manchmal ein wenig für Verwirrung sorgen. Belassen Sie es also hier

nicht nur bei der Definition, sondern erklären Sie, warum dieser Wert so wichtig ist.

Wurden die Kriterien erst einmal benannt, erklärt und verstanden, so ist es für den Studierenden nicht mehr allzu schwer, die Validität dieser Studie zu evaluieren. So würde der Studierende trotz des nicht verblindeten Studiendesigns (i.m.-Injektion versus Tabletten) und des vielleicht ein wenig zu langen Follow-up-Zeitraumes (Reinfektion innerhalb von 9 Monaten?) der Studie eine gute Validität zusprechen.

Sicherlich würden dem EbM-Experten bei detaillierter Analyse der besagten Studie weitere negative Punkte auffallen. Vergessen Sie jedoch bitte nicht, dass es sich bei den Studierenden in der Regel um EbM-Anfänger handelt. Die Tatsache, dass die Therapiestudien überhaupt nach Validitätskriterien überprüft werden, ist als großer Erfolg zu werten.

Ist die Validität der Studie erst einmal geklärt, geht es nun um die Relevanz. Wie oben bereits beschrieben, wird dieser Teil von den Studierenden aufgrund der Biometrie und der mit ihr verbundenen „Trockenheit“ oft gemieden. Bringen Sie also den Studierenden grundsätzlich nur die Werte und Definitionen bei, die diese später im klinischen Alltag auch verwenden würden. Einen z-Test herzuleiten, würde jeden Studierenden – und sicherlich auch den ein oder anderen Kliniker – verwirren. Um Therapiestudien bewerten zu können, reichen hierbei durchaus die Kenngrößen Control event rate (CER), Experimental event rate (EER), Absolute Risiko-Reduktion (ARR), Relative Risiko-Reduktion (RRR) und Number needed to treat (NNT) [4].

Alle Werte bis auf die NNT werden erfahrungsgemäß sehr leicht verstanden. Für die Vermittlung der NNT sollten Sie sich jedoch etwas Zeit nehmen und vielleicht an einem Beispielsatz verdeutlichen (z.B. „Behandle X Patienten mit der neuen Methode, um ein negatives Ereignis mehr zu verhindern.“).

Auch sollte der Unterschied von ARR zu RRR verdeutlicht werden. Nützlich für dieses Unterfangen können auch die Broschüren mancher Pharmafirmen sein.

Bevor es jedoch an die Kennzahlen geht, sollte der Studierende in der Lage sein, die zur Berechnung notwendigen Werte der Studie zu entnehmen. Gemeint ist hiermit die Definition des negativen Ereignisses. Selbst das Verständnis der fünf Kenngrößen würde bei der Evaluation der Relevanz nichts nützen, wenn der Studierende nicht erkennt, womit er die Kenngrößen berechnen soll. Erinnern Sie ihn dabei an die zuvor definierte vierteilige Frage. Die definierte Zielsetzung bei dieser Frage sollte den Studierenden daran erinnern, nach welchem negativen Ereignis er zu suchen hat.

Der Studierende entscheidet sich bei der Definition des negativen Ereignisses für die kumulative Heilungsrate der gesamten Studienpopulation nach neun Monaten, erkennbar aus Tabelle 2 der betreffenden Studie. Auch dass er von dem positiven Ereignis der „Heilungsrate“ auf die Rate des negativen Ereignisses „zurückrechnen“ muss, erkennt der Studierende im Allgemeinen selbst.

So erhält nach kurzer Rechenpause der Studierende eine CER von 5,0%, eine EER von 2,3%, eine ARR von 2,7% (RRR=54%) und eine NNT von 38.

Nun gilt es, die gewonnenen Erkenntnisse auf das Fallbeispiel übertragen zu können. Hierbei kommt es darauf an, wie viel klinische Erfahrung der Studierende bereits hat. So könnte er unter Verwendung der ARR und NNT sowohl zugunsten des Azithromycin („Wieso eine Injektion geben, wenn es eine Tablette auch tut?“) als auch zugunsten des Penicillins („Wieso das Risiko der Resistenz eingehen, schließlich wirkt Penicillin doch?“) argumentieren [5]. Sie können natürlich seine Sichtweisen durch Ihre eigene Sicht des Sachverhaltes erweitern (z.B. „... bisher keine bekannten Resistenzen von *Treponema pallidum* gegenüber Penicillin ...“). Seien Sie

dabei jedoch nicht allzu streng und überlassen Sie dem Studierenden ein wenig Freiraum für die Übertragung auf das Fallbeispiel – vorausgesetzt natürlich, dass seine Argumentationen nicht allzu weit von der Realität abweichen.

Nun ist es soweit – der Studierende hat eine therapeutische Fragestellung gelöst, indem er sich die Information selbst eingeholt und selbst bewertet hat. Sie können sich vielleicht vorstellen, was für ein Erfolgserlebnis dies für den Studierenden sein kann. Er hat dadurch nicht nur eine zusätzliche medizinische Information erhalten, sondern ist von nun an in der Lage, Recherchen selbst durchzuführen und gefundene Evidenzen grundlegend zu bewerten.

39.4 Sie können 12 Tipps für einen gelungeneren Unterricht aufzählen

In den Beispielszenarien ist Ihr Lernumfeld jedes Mal ein anderes. Durch unsere Erfahrung in der Ausbildung und den regen Austausch mit Studierenden konnten wir einige allgemeine Tipps zusammenstellen, die einem Unterricht in Evidenzbasierter Medizin gut tun können.

- ▲ **Holen Sie die Studierenden da ab, wo sie stehen.** Es ist Ihre Aufgabe zu erkennen, wo die Studierenden anfangs mit ihrem Wissen stehen. Holen Sie sie genau dort ab. Es ist nämlich Ihr Erfolg, wenn sie nach der Ausbildung dort ankommen, wo Sie sie von Beginn an haben wollten.
- ▲ **Formulieren Sie Lernziele.** Sagen Sie am Anfang so genau wie möglich, was Sie von den Studierenden erwarten. Dann können die Studierenden darauf hinarbeiten und leichter zwischen wichtigen, interessanten oder anekdotischen Anteilen unterscheiden. Achten Sie darauf, für Ihre Studierenden relevante Lernziele zu erstellen. Die Studierenden müssen den Nutzen des vermittelten Wissens erkennen. Um Lernziele gut zu formulieren bedarf es ein wenig Übung. Eine Hilfe ist die SMARTe Erstellung von Lernzielen. Das heißt: Lernziele sollen
 - Spezifisch,
 - Messbar,
 - Attraktiv,
 - Relevant und
 - Terminiert sein.
 Es sollte genau geklärt werden was in welchem Umfang bis wann erlernt werden soll und warum es wichtig ist. Beispielsweise könnte ein Lernziel lauten: Am Ende dieses Buchkapitels soll der Leser den Unterschied zwischen formativer und summativer Evaluation erklären können, um Unterrichtsevaluationen gezielter einzusetzen.
- ▲ **Kontrollieren Sie Ihre Lernziele.** Ihre Lernziele sind wichtig und Sie wissen, warum. Deshalb nehmen Sie diese Ziele ernst und fragen Sie diese Ziele ab. Fragen Sie die Studierenden außerhalb des aktuellen Zusammenhangs, um die Information in verschiedenen Kontexten zu wiederholen.
- ▲ **Gehen Sie davon aus, dass der Studierende lernen will.** Auch wenn dies nicht stets zutrifft, können Sie dies voraussetzen.
- ▲ **Spekulieren Sie** – nicht an der Börse, sondern mit den Studierenden. In unseren Seminaren haben wir die Erfahrung gemacht, dass viele Studierende Angst davor haben, etwas Falsches zu sagen. Deshalb wird häufig gar nichts gesagt. Ermutigen Sie die Studierenden zu mehr Spekulation auf der Grundlage des vorhandenen Wissens, und fragen Sie gezielt einzelne Personen. Nur so erfahren Sie etwas über den tatsächlichen Wissensstand und binden die Aufmerksamkeit.
- ▲ **Hören Sie zu.** So nehmen Sie den Studierenden ernst. Hören Sie in Ruhe zu, denken darüber nach und formulieren dann

eine Antwort. Ein Kommentar oder eine Frage bedeutet aktive Beteiligung. Machen Sie klar, dass dies ernst genommen wird.

- ▲ **Beteiligen Sie alle.** Häufig werden Fragen oder Kommentare nicht verständlich formuliert. Wiederholen Sie sie, fragen Sie andere Studierende nach ihrer Meinung. Dadurch merken die Studierenden, dass andere das gleiche Problem hatten, aber einfach nicht gefragt haben. Resümieren Sie wichtige Zusammenhänge, denn keiner schafft es, ununterbrochen aufmerksam zu sein. Sie ermöglichen dadurch den Wiedereinstieg in die Thematik.
- ▲ **Warten Sie eine Antwort ab.** Gelegentlich kommt auf eine Frage keine Antwort. Dies kann 3 Ursachen haben: Entweder die Frage war zu leicht oder zu schwer oder es hat niemand aufgepasst. Machen Sie nicht den Fehler, die Frage selbst zu beantworten. Warten Sie einfach ab. Die entstehende Stille müssen nicht nur Sie ertragen, sondern auch die Studierenden. Seien Sie standhaft, sonst gehen Ihre Teilnehmer davon aus, dass die nächsten Fragen auch von Ihnen selbst beantwortet werden ... nach ausdauerndem Schweigen.
- ▲ **Lassen Sie die Studierenden ran.** Erlauben Sie Ihnen, die Schritte der Evidenzbasierten Medizin selbst durchzuführen – zum Beispiel an einem konkreten Fallbeispiel. So erzeugen Sie Erfolgserlebnisse, die Studierenden begreifen die Abfolge und den Inhalt der Evidenzbasierten Medizin und erfahren die praktische Umsetzbarkeit.
- ▲ **Lassen Sie die Studierenden untereinander lehren.** Geben Sie einen Rahmen für den gegenseitigen Austausch zur Festigung des Wissens. Sie können dies z.B. im Rahmen eines ersten Journal Clubs ermöglichen oder Kleingruppenarbeiten fördern.
- ▲ **Seien Sie Vorbild.** Zeigen Sie, dass Evidenzbasierte Medizin im Alltag umsetz-

bar ist, indem Sie sie umsetzen. Bringen Sie Beispiele für den Nutzen in Ihrem Alltag und verknüpfen dies mit den Zielen der Studierenden.

- ▲ **Übertragen Sie das Erlernte.** Nehmen Sie verschiedenste Situationen, in denen die Grundlagen der Evidenzbasierten Medizin von Nutzen sind, und sprechen Sie darüber. So könnten Sie gesundheitspolitische Entscheidungen durchleuchten, Neuerungen in ganz verschiedenen medizinischen Bereichen erklären oder auch über die Planung einer Promotionsarbeit sprechen.

39.5 Sie können die Eigenschaften von formativen und summativen Evaluationen definieren

Der Lernerfolg und der Lehrerfolg können auf zwei Arten kontrolliert werden. Die so genannte **formative Evaluation** dient der weiteren Verbesserung des Lernens oder Lehrens. Die so genannte **summative Evaluation**, die eigentliche Prüfung, dient der Selektion der Lernenden, aber auch möglicherweise der Lehrenden. Formative Evaluationen sind in den Prozess des Lernens oder Lehrens integriert, während summative Evaluationen Prüfungen bezeichnen, die ein abschließendes Ergebnis darstellen [6].

Die formative Evaluation. Sie dient der Evaluation der Verbesserung des Lernerfolges, hat keine Konsequenz für die Studierenden und gibt nur die Möglichkeit, Schwächen und Stärken einzuschätzen. So werden z.B. Probeexamina durchgeführt.

Dient die Evaluation der Verbesserung der Lehre, kann hier die Lehrkraft die Stärken und Schwächen des Unterrichts einschätzen. Dies kann zum Beispiel über subjektive Evaluationen (Feedback) geschehen.

Die summative Evaluation (= Prüfung). Die wohlbekanntesten Prüfungen der Studierenden dienen der Selektion, d.h. der Über-

prüfung des Erreichens der gestellten Anforderungen. Diese kann im Sinne von Noten zusätzlich abgestuft sein, je nachdem, wie gut die Anforderungen erfüllt wurden.

Selektion aufgrund der Überprüfung des Lehrerfolges tritt dann ein, wenn entschieden wird, ob ein bestimmtes Curriculum eingeführt wird oder bestehen bleibt. Dazu werden zum Beispiel fachübergreifende, meist subjektive Evaluationen eingesetzt.

Prüfungen dienen 3 Zwecken. Sie stellen die **Anforderungen**, welche der Studierende erbringen soll, sie messen den **Lernerfolg** des Studierenden und den **Lehrerfolg** des Tutors. Im letzten Fall werden sie häufiger Evaluationen genannt.

Die Anforderung bestimmt die Lernleistung. Unserer Meinung nach wissen Studierende sehr genau, wie viel (Lern-)Aufwand nötig ist, um eine Klausur oder einen Test zu bestehen. Nicht umsonst gibt es gut gehetzte Klausursammlungen und Prüfungsprotokolle an den meisten Universitäten. Eine bewusste Prüfung kann die Studierenden dazu bringen, das Richtige zu lernen. Wenn Sie also wollen, dass die Studierenden etwas sicher (er-)lernen, prüfen Sie es.

Sie bestimmen den Zweck einer Evaluation oder Prüfung. Selbst keine Prüfung ist eine Entscheidung. Dann setzen Sie die Anforderungen allerdings auf null. Die Studierenden kostet die Ausbildung nichts an Lernaufwand. Höchstens freiwillig können Sie etwas geben. Und was nichts kostet, kann bekanntlich auch nichts taugen. Übersetzt könnte es vielleicht lauten: Was nicht geprüft wird, ist nicht wichtig. Die Frage ist: Was wollen Sie erreichen. Ist Ihr Ziel die Selektion, die Überprüfung Ihres Konzeptes oder das Vermitteln von Wissen? Daran können Sie Ihre Form der Prüfung festmachen. Um Unterricht und Prüfungen aufeinander abzustimmen, empfiehlt es sich, vorher die Anforderungen an die Studierenden festzulegen. Erstellen Sie Lernziele.

39.6 Stand der Ausbildung und Ausblick

Die Evidenzbasierte Medizin wurde ursprünglich als neuer Weg in der praxisorientierten medizinischen Lehre entwickelt. Die Evidence-based Medicine Working Group um Gordon Guyatt veröffentlichte 1992 unter dem Begriff der „Evidence-based Medicine“ ein Konzept, um Medizinern die Möglichkeit zu geben, sich eigenständig, hochaktuell und qualitativ hochwertig weiterzubilden [7].

Dieses Konzept wurde vor allem in Nordamerika rasch aufgegriffen und in dortige Curricula zur Ausbildung integriert. Ende des vergangenen Jahrzehnts wurden erstmals auch in Deutschland erste explizite EbM-Kurse für Studierende durchgeführt. Seitdem werden regelmäßig in Deutschland Freiwilligenkurse, an denen auch Studierende teilnehmen können, durchgeführt. Im Jahre 2002 wurde der Fachbereich „EbM im Studium“ des Deutschen Netzwerks Evidenzbasierte Medizin e.V. für den deutschsprachigen Raum gegründet, um eine Plattform für EbM-Lehrende im Bereich der Ausbildung zu etablieren. EbM-Aktive unter anderem aus Witten-Herdecke, Lübeck, Ulm, Frankfurt und vielen anderen Städten sind Mitglieder dieses Fachbereiches und verfolgen damit gemeinsam das Ziel der Implementierung von EbM im Medizinstudium. Es ist erfreulicherweise an einigen Universitäten bereits gelungen, EbM auf verschiedene Arten in das Studium zu integrieren. Neben studienbegleitenden EbM-Programmen wie in Witten-Herdecke werden andernorts Teile der EbM in bestehende Curricula (Lübeck) integriert oder auch explizite, scheinpflichtige EbM-Seminare durchgeführt (Frankfurt). Neue Reformstudiengänge integrieren von vornherein Inhalte der EbM in ihre Curricula (Aachen, Berlin).

Damit nutzen diese Universitäten gemeinsam die Möglichkeiten der EbM, eine

fundierte wissenschaftliche Ausbildung mit einer eigenverantwortlichen und selbstständigen Berufsausübung und der Fähigkeit zur ständigen Fortbildung zu verbinden.

Anhand der oben genannten Beispiele ist zu ersehen, dass die Durchführung der evidenzbasierten Ausbildung machbar ist. Jedoch ist die Implementierung von EbM in bestehende Curricula nicht immer einfach und komplikationslos. Neben Schwierigkeiten mit (bestehenden) Rahmenbedingungen, zum Beispiel durch fehlende Infrastrukturen, ist die größte Hürde bei der Implementierung sicherlich die Initialisierung von EbM-Aktivitäten an einem Standort. Sowohl die anfängliche Organisationsarbeit als auch die vielleicht verbesserungsfähige Qualifikation für Lehre in EbM machen es Engagierten oft schwer, eine Ausbildung in EbM im curricularen Umfeld einzurichten. Zur mittelfristigen Etablierung einer kontinuierlichen, curricularen Ausbildung von Studierenden in EbM sind ausreichend Lehrende an einem Standort notwendig, die diese Aufgabe qualifiziert übernehmen können.

Um die Lehre der Evidenzbasierten Medizin zu fördern, wird mit Unterstützung der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), des DNEbM und der Gesellschaft für Medizinische Ausbildung (GMA) ein Multiplikatoren-Programm angeboten. Dieses richtet sich an Lehrende, die ihre Studierenden in Evidenzbasierter Medizin ausbilden wollen. Es beinhaltet tiefer gehende Ausbildungsschritte über einen Grundkurs hinaus sowie ein Lehr- und Moderatorentaining (nähere Informationen dazu unter www.train-the-teacher.de).

Literatur

- [1] Approbationsordnung für Ärzte. Bundesgesetzblatt. (2002), 1, 2405–2435
- [2] Sackett et al. (2000) Evidence-Based Medicine. How to Practice and Teach EbM. 2nd edition. Churchill Livingstone. London
- [3] Riedner G et al., Single-dose azithromycin versus penicillin G benzathine for the treatment of early syphilis. *N Engl J Med.* (2005), 353, 1236–1244
- [4] Antes G et al. (2005) Curriculum EbM. Ärztlicher Fortbildungskatalog Evidenzbasierte Medizin., Berlin
- [5] Lukehart SA et al., Macrolide resistance in *Treponema pallidum* in the United States and Ireland. *N Engl J Med.* (2004), 351, 154–158
- [6] Scriven M (1991) Evaluation thesaurus. Sage Publ., Newbury Park
- [7] Evidence-Based Medicine Working Group, Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA.* (1992), 268, 2420–2425

40 Weiter- und Fortbildung

Norbert Donner-Banzhoff, Olaf Weingart, Gerd Burmester, Susanne Weinbrenner, Hans Wille, Regina Kunz

40.1 EbM: Von der Etablierung zur Routine

In der erste Dekade der evidenzbasierten Medizin in Deutschland ging es um die Etablierung eines neuen Paradigmas. Kliniker (vielfach mit Erfahrungen in Kanada oder Großbritannien), Methodiker und Angehörige von Spitzenorganisationen und Instituten der Selbstverwaltung trieben diesen Prozess voran. Als Medien dienten Zeitschriften (z.B. Serie in der Münchner Medizinischen Wochenschrift/MMW, die Zeitschrift für ärztliche Fortbildung und Qualität im Gesundheitswesen (ZaeFQ – <http://www.zaefq.de>)), Lehrbücher und vor allem der „klassische“ EbM-Kurs. Letzterer sprach einen überregionalen Kreis von Adressaten (meist Ärzten) an, die sich in ihrer Region über ihre engeren klinischen Aufgaben hinaus in Lehre, Forschung oder Qualitätsmanagement engagierten. Diese Veranstaltungsform wurde im EbM-Curriculum (siehe unten) formalisiert.

Inzwischen ist der Bezug auf die Evidenz wissenschaftlicher Studien allgemein etabliert; darüber sollten Diskussionen über die Gewichtung (z.B. von Leitlinien) nicht hinwegtäuschen. Grundkenntnisse sind gemäß der neuen Approbationsordnung bereits im Medizinstudium zu vermitteln (s. Kap. 39). Vor diesem Hintergrund müssen wir die Rolle der EbM-Dimension in Weiter- und Fortbildung neu definieren.

Von allen Gesundheitsprofessionen wird heute ein kontinuierlicher Lern- und Anpassungsprozess verlangt. Die Vielfalt und der Umfang unserer Informationsquellen, deren oft widersprüchliche Aussagen und die Be-

einflussung durch Interessen machen ein kritisches Informationsmanagement zur Überlebens-Qualifikation. Jeder Arzt ist irgendwann mit neuen und/oder ungewöhnlichen Problemen konfrontiert, die er durch aktive Recherche im weiteren Sinne lösen muss; dazu kommen die kritischen Fragen informierter Patienten, die nach fundierten Antworten zu Prognose sowie diagnostischer und therapeutischer Wirksamkeit verlangen. EbM umschreibt damit eine generische Kompetenz, die für eine medizinische Tätigkeit ebenso wichtig ist wie Grundkenntnisse in Anatomie oder Biochemie.

Trotz dieser Erfordernisse ist unser Bewusstsein dafür gewachsen, dass nicht alle Ärzte die Kompetenz aufweisen müssen, auf die der klassische EbM-Kurs zielt, z.B. eine englischsprachige Originalarbeit kritisch zu diskutieren (sog. „appraising mode“ [1]). Alle Ärzte sollten allerdings den „searching mode“ beherrschen, um aufbereitete – vorzugsweise evidenzbasierte – Quellen aktiv zu nutzen; nur so werden sie Antworten auf eigene und patientengenerierte Fragen finden.

EbM in der Weiter- und Fortbildung zielt also auf ein differenziertes Kompetenzspektrum innerhalb der Gesundheitsprofessionen. Dabei gehen wir davon aus, dass für folgende Gruppen EbM-Fertigkeiten auf dem Niveau eines einfachen Searching-Mode nicht ausreichen:

- ▲ Entwickler von Leitlinien und Klinischen Pfaden,
- ▲ Moderatoren von Qualitätszirkeln,
- ▲ Betreuer (Dozenten, Moderatoren) von Fortbildungs-Veranstaltungen.

Gemeinsam ist diesen Gruppen, dass es sich um lokale Meinungsbildner handelt, die jeweils für eine größere Zahl von Kollegen eine „Informationsverantwortung“ übernehmen. Darüber hinaus werden sich natürlich in allen Arbeitsbereichen Kolleginnen und Kollegen von der EbM angesprochen fühlen, so dass sie auch ohne die genannten Aufgaben ihre EbM-Fertigkeiten entwickeln wollen.

Welche inhaltliche Ausrichtung und welche didaktischen Formate können helfen, das oben skizzierte differenzierte Kompetenzspektrum zu sichern? Im Folgenden diskutieren wir drei Beispiele aus den inzwischen zahlreichen Möglichkeiten.

40.2 Didaktische Formen

40.2.1 Fallbeispiel 1: Der „klassische“ EbM-Kurs

In seinem Aufbau und seiner Durchführung orientiert sich der typische EbM-Kurs an den Workshops der McMaster University in Hamilton, Kanada, dem Gründungsort der evidenzbasierten Medizin. Diese Universität ist auch die Geburtsstätte des problemorientierten Lernens in der studentischen Ausbildung. Dessen Elemente finden sich deshalb auch im klassischen EbM-Kurskonzept wieder: intensive Kleingruppen-Arbeit, moderiert durch einen Tutor, wenig Frontalunterricht, Orientierung an einem ganz konkreten Patientenproblem, das mit dem Gelernten gelöst werden soll.

Die fünf Schritte der evidenzbasierten Medizin dienen als Richtschnur:

- ▲ Fragestellung,
- ▲ Recherche einer Studie,
- ▲ Kritische Beurteilung,
- ▲ Übertragung auf den einzelnen Patienten,
- ▲ kritische Selbsteinschätzung [1].

Schlüsselfragen – auch diese gemäß den kanadischen Vorbildern standardisiert – die

nen dem Lernenden als „Leitplanken“ für die Studienbewertung.

Der Umfang der Workshops variiert: während der Klassiker über 5 Tage stattfindet und ein breites Spektrum an klinischen Fragestellungen und den zugehörigen Studien abdeckt (Therapie, Diagnose, Prognose, Nebenwirkungen und Ätiologie, Meta-Analyse, Leitlinien, Entscheidungsanalyse, Gesundheitsökonomie), werden in Deutschland meist schlankere Varianten von ein bis zwei Tagen Dauer angeboten, die sich auf Fragen zu Therapie und Diagnostik, Meta-Analysen und Leitlinien konzentrieren. Kleingruppensitzungen werden häufig ergänzt durch Vorträge, praktische Übungen zur Literaturrecherche oder biometrischen Fragestunden.

Der klassische Kurs wendete sich ursprünglich an alle Ärzte vom Berufsanfänger bis zum leitenden Oberarzt oder Chefarzt, wie auch an Ärzte aus der Systemebene (Krankenkassen, ärztliche Selbstverwaltung, Industrie usw.) Später kamen Kurse für medizinische Laien dazu: Verbraucher, Patientenvertreter, Journalisten (s. Kap. 41).

Die Kurse finden regional oder überregional statt, meist außerhalb des Tätigkeitsbereichs der Teilnehmer, wodurch konzentriertes Lernen in einer ungestörten Atmosphäre ermöglicht wird, und integrieren meist Kollegen aus unterschiedlichen medizinischen und chirurgischen Disziplinen. Fokussierung auf ein bestimmtes Fachgebiet, wie z.B. Pädiatrie, Gynäkologie oder Onkologie, sieht man seltener. Inzwischen dürften in Deutschland mehr als zehntausend Personen an einem solchen Kurs teilgenommen haben.

Die Teilnehmer haben diese Kurse nicht nur hervorragend evaluiert. Tatsächlich ist im Rahmen der Berliner Grundkurse auch ein relevanter Zuwachs an Wissen und Fertigkeiten empirisch nachgewiesen worden [2]. Dazu kommt der motivierende Effekt, sich mit der EbM weiter auseinanderzusetzen, die Stärken und Grenzen der Methode auszutesten.

40.2.2 Fallbeispiel 2: Journal Club

Das Bremer Institut für Klinische Pharmakologie veranstaltet seit 1998 regelmäßig Kurse zur evidenzbasierten Medizin. Bereits in den ersten Kursen erkannten Teilnehmer die fehlende Kontinuität im Lesen von Studien als relevantes Problem. Aus diesem Grund etablierte sich parallel zu den EbM-Kursen ein Journal Club (JC). Er bietet den Kursabsolventen und anderen EbM-Interessierten die Möglichkeit, das erworbene Wissen regelmäßig anzuwenden und zu vertiefen.

Seit 1998 findet der JC einmal monatlich im Bremer Institut statt. 10–20 Teilnehmer aus dem stationären, niedergelassenen, aber auch pharmazeutischen Bereich nennen interessierende Themen oder Studien. Themenschwerpunkte sind aktuelle Publikationen überwiegend aus dem allgemeinmedizinischen oder internistischen Bereich. Eine Woche vor dem jeweiligen Termin versenden die für den Journal Club Verantwortlichen des Instituts die Studie als PDF-Anhang gemeinsam mit der Einladungsmail.

Zu Beginn wird die Studie kurz durch einen Teilnehmer vorgestellt. Unter Moderation durch einen Mitarbeiter des Instituts werden Hintergrund und Relevanz der Fragestellung erörtert. Hierbei erweist sich die Interdisziplinarität als großer Vorteil. Durch die verschiedenen Perspektiven (z.B. Niedergelassener vs. Klinikarzt) werden die Diskussionen untereinander sehr differenziert geführt. Bei den Teilnehmern besteht mittlerweile ein respektables Rüstzeug im kritischen Lesen von Studien, sodass die Mitarbeiter des Instituts überwiegend eine moderierende Funktion haben. Trotzdem hat sich das Vorgehen anhand einer Checkliste für Inhalt und Methodik bewährt. Das selbstständige Erarbeiten der jeweiligen Publikation wird einer vorgetragenen Präsentation vorgezogen. Hilfreich kann jedoch die ergänzende Darstellung zentraler Ergebnistabellen oder -kurven sein. Da an diesem JC überwiegend

praktisch tätige Ärzte teilnehmen, sollte am Ende eines JC immer ein praktisches Fazit – als „Take home message“ – stehen.

In Bremen hat sich der JC als fester Bestandteil der Fortbildungsaktivitäten im Bereich EbM etabliert. Hilfreich ist hierbei der ständig aktualisierte E-Mail-Verteiler, der sich hauptsächlich aus früheren EbM-Kursteilnehmern zusammensetzt und zur Zeit über 100 Interessenten umfasst. Bei 15 EbM-Kursen in den letzten Jahren sind die Bremer Ärzte intensiv mit der EbM in Kontakt gekommen. Unter den o.g. Umständen scheint somit zumindest in den vergangenen 8 Jahren die „kritische Masse“ für einen regelmäßigen JC erreicht worden zu sein.

40.2.3 Fallbeispiel 3: PERLEN

„Patientenorientierte, evidenzbasierte RECHERCHE – Lernen, Entwickeln, Nutzen“ stellt eine Lernhilfe für Hausärzte dar, die von der Abteilung für Allgemeinmedizin der Universität Marburg entwickelt worden ist. PERLEN zielt auf Kompetenzen im Searching-Mode: Formulieren von „suchtauglichen“ Fragen, die Recherche in aufbereiten, aber spezifischen Quellen und die kritische Bewertung von Aussagen. Als Neuerung sind Such-Heuristiken vorgesehen, die gerade auch einen raschen Abbruch der Suche ermöglichen, ohne zu viel an Sensitivität zu opfern. Eine begrenzte Palette von Quellen wird mit einem Bewertungssystem nach Aufbereitungs-Grad, Umfang, Suchmöglichkeiten und Bias dargestellt. Wissenschaftliche Datenbanken wie MEDLINE und EMBASE werden nicht detailliert behandelt, dafür aber verbraucherfreundliche, d.h. aufbereitete Quellen wie z.B. infomed-screen, das Arzneimitteltelegramm oder Clinical Evidence. Zur Verfügung stehen:

- ▲ ein Veranstaltungskonzept (Übungen an Mehrplatz-PC-Anlagen), das z.B. in Qualitätszirkeln angeboten werden kann;

- ▲ eine Broschüre zum Selbststudium;
- ▲ eine CD-ROM mit Hyperlinks zu einschlägigen Quellen mit zusätzlichen Erläuterungen und Ressourcen;
- ▲ ein Manual für Betreuer bzw. Moderatoren.

Zwar werden wenige definierte Beurteilungskriterien für klinische Studien behandelt (z.B. Studientypen, klinisch relevante Zielkriterien, Wirksamkeitsmaße); wissenschaftliche Originalpublikationen selbst bzw. deren Beurteilung sind jedoch bewusst ausgeklammert. Als wesentliche Stoßrichtung hat sich PERLEN vorgenommen, Ärzte dazu zu motivieren, ihre eigenen Fragen aus der Versorgung ernst zu nehmen, klar zu formulieren und weiter zu verfolgen. In einer ersten Pilotstudie fand PERLEN eine positive Resonanz bei einer Gruppe von Hausärzten. Die Evaluation im Rahmen einer randomisierten kontrollierten Studie dauerte bei Redaktionsschluss noch an.

40.2.4 Weitere Möglichkeiten

Von Veranstaltungen, die auf eine **allgemeine Kompetenz** der Teilnehmer zielen, hat sich das Spektrum in Richtung von Formaten erweitert, die sich stärker an evidenzbasierten Lösungen für **Probleme aus der Versorgungspraxis** orientieren. Letztere legen besonderen Wert auf die Integration von EbM in die alltägliche Routine und zielen auf Einstellungen und konkrete Verhaltensänderungen von Ärzten. So wurden beispielsweise Praxisbesuche eingesetzt (Outreach, Academic Detailing) [3] und/oder an reale Patientenprobleme angeknüpft [4]. Eine weitere Möglichkeit der Integration bieten EbM-gestützte persönliche Auskunftssysteme, die allerdings nur sehr vereinzelt dauerhaft zur Verfügung stehen (siehe beispielsweise <http://www.attract.wales.nhs.uk> [5]). Sehr wirksam ist die Vermittlung von EbM-Fertigkeiten für bestimmte Zielgruppen mit einem spezifischen Lernbedarf, z.B. Leitlinien-Autoren.

40.3 Das Curriculum Evidenzbasierte Medizin der Bundesärztekammer und des DNEbM e. V.

Vor dem Hintergrund der politischen Diskussion um die Qualität der Versorgung beschlossen die Träger des Ärztlichen Zentrums für Qualität (ÄZQ) – die BÄK und die KBV – im Jahr 2000, einen Lehr- und Lernzielkatalog für die Ärztliche Fortbildung im Bereich der EbM (in Analogie zu den Curricula für Fortbildung der Bundesärztekammer) zu entwickeln.

Dieses „Curriculum Evidenzbasierte Medizin“ wurde durch einen interdisziplinären Expertenkreis des ÄZQ in Zusammenarbeit mit dem Deutschen Netzwerk Evidenzbasierte Medizin (DNEbM) e. V. erarbeitet und im Jahre 2002 in der 1. Auflage fertig gestellt. Mit der landesweiten Implementierung des Curriculums EbM sollen folgende Ziele erreicht werden:

- ▲ Ärzten, Patienten und der Öffentlichkeit valide und verständliche Information über diagnostische, therapeutische und präventive Maßnahmen in der Gesundheitsversorgung zu vermitteln.
- ▲ Die Begründung ärztlicher Maßnahmen auf wissenschaftlicher Grundlage zu reflektieren und zu verbessern.
- ▲ Die Nutzung von EbM als Werkzeug eines internen QM-Systems zu fördern.
- ▲ Diskurse über Prioritäten, Ressourcen und Qualität im Gesundheitswesen besser zu fundieren.
- ▲ Die EbM in die Aus-, Weiter- und Fortbildung der Ärzte einzubinden (<http://www.dnebm.de>).

Nach Abschluss einer Test- und Evaluationsphase erfolgte eine Überarbeitung. Die 2. Auflage von 2005 entspricht den Vorgaben des Deutschen Senats für ärztliche Fortbildung der Bundesärztekammer (BÄK) und ist von diesem als ärztlicher Fortbildungskatalog der BÄK für evidenzbasierte Medizin anerkannt worden.

Das Curriculum gibt die methodische Grundlage, evidenzbasierte Medizin systematisch und strukturiert zu lehren und zu lernen. Es gliedert sich in zwei Teile (Grundkurse und Kurse für Fortgeschrittene), die aufeinander aufbauen und die Möglichkeit bieten, unterschiedliche Schwerpunkte zu setzen.

Die **Grundkurse** dienen der Einführung in die Methode EbM. Hier ist die Diskussion einer Studie zu „Therapie/Intervention/Prävention“ obligat, weitere Themen bzw. Kompetenzen können flexibel eingebaut werden. Dazu kommen übergreifende Themen wie Statistik, Informationsrecherche und Patienteninformationen.

Die **Kurse für Fortgeschrittene** sollen eine gezielte Vertiefung in weitergehende Konzepte der EbM ermöglichen:

- ▲ Systematische Übersichtsarbeiten,
- ▲ Evidenzbasierte Leitlinien,
- ▲ Evidenzbasierte Patienteninformationen,
- ▲ Lehre und Verbreitung – „Train-the-trainer“.

Das Curriculum EbM soll primär Veranstaltern als Grundlage für die inhaltliche und organisatorische Planung ihrer Bildungsangebote dienen und ist Grundlage der seit 2006 vom Deutschen Netzwerk EbM angebotenen Zertifizierung von Kursen (siehe <http://www.ebm-netzwerk.de>).

40.4 EbM-International: Leonardo da Vinci

In einigen europäischen Ländern sind EbM-Kenntnisse und -Fertigkeiten bereits optionaler oder obligater Bestandteil der Facharzt-Weiterbildung und der ärztlichen Fortbildung. Da die zunehmende Mobilität auf dem Arbeitsmarkt Europa auch eine Vereinheitlichung von Qualifikationen erfordert, fördert das **Leonardo da Vinci-Programm** der Europäischen Union Programme zur berufsbegeleitenden Weiterqualifizierung, um diese



Abb. 40.1: Logo des EUebm-Projekt

Flexibilisierung des Arbeitsmarktes innerhalb Europas zu erleichtern.

Bei dem integrierten EbM-Curriculum des **EUebm-Projekts** handelt es sich um eine Weiterentwicklung des EbM-Lehrens und -Lernens im unmittelbaren Kontext der Patientenversorgung, bei dem lokale, regionale, nationale und europäische Erfahrungen und Kenntnisse zusammengebracht und gebündelt werden. Das Curriculum soll Ärzten zu dem Wissen und den notwendigen Fertigkeiten verhelfen, mit denen sie zeitnah die beste vorhandene Evidenz finden können, um gemeinsam mit dem Patienten informierte Entscheidungen über das weitere diagnostische und therapeutische Vorgehen zu treffen. Dabei sollen nationale Besonderheiten der Weiter- und Fortbildung berücksichtigt werden.

Für das EUebm-Projekt entwickeln die Projektpartner in einem Netzwerk aus neun europäischen Ländern (Großbritannien, Deutschland, Österreich, Niederlande, Polen, Spanien und Ungarn, Italien und der Schweiz) ein gemeinsames Verständnis über Inhalt und Umfang des Modellprojekts, mit besonderem Augenmerk auf den Möglichkeiten von Lern- und Lehrsituationen im beruflichen Alltag (Integrated Teaching) [4]. Dazu gehören spezifische Lernzieldefinitionen, angepasste Lerntools und eine Internet-Plattform für Online-Lernmodule. Die nationalen Partner erhalten die Möglichkeit, das gemeinsame Kern-Curriculum entsprechend den nationalen Gegebenheiten weiter auszubauen und den Anforderungen der ärztlichen Weiter- und Fortbildung anzupassen.

Das Kern-Curriculum mit den Umsetzungsmodulen wird in den beteiligten Ländern pilotiert und vor der Publikation entsprechend überarbeitet. Zunächst werden

nationale Akkreditierungen gesichert, die in eine übergreifende europäische Anerkennung münden sollen. Nach dem Abschluss des Projekts können lokale und nationale Initiativen aus ganz Europa auf die Ergebnisse dieses Projekts zurückgreifen, sei dies in Bezug auf die Planung von Curricula, didaktische Ideen oder erprobte Materialien.

40.5 Die EbM-Zukunft

Wohin wird uns die nächste EbM-Dekade bringen, welche Wege wollen wir einschlagen? Nach der reformierten Approbationsordnung sollen künftige Hochschulabsolventen mit den Grundbegriffen der EbM besser vertraut sein als ihre Vorgänger. Es wird ihnen damit aber genauso gehen wie mit den direkt klinischen Lerninhalten: sie werden im Rahmen der Weiterbildung den Praxisbezug noch entwickeln müssen.

Der wichtigste Motivator in diesem Prozess sind identifikationsfähige Vorbilder, d.h. „gestandene“ Kliniker und Praktiker, die eine Synthese von individueller Erfahrung und wissenschaftlicher Verankerung im Alltag umsetzen und vorleben. Außerdem müssen spezifische Quellen und deren Nutzung behandelt werden, z.B. Zeitschriften mit Evidenzbezug oder auf das jeweilige Fach zugeschnittene Datenbanken.

Unserer Meinung nach sind begleitende Kurse im Rahmen der Weiterbildung hierzu das beste Mittel; dies zeigen auch die internationalen Erfahrungen aus strukturierten Weiterbildungsprogrammen (z.B. in Großbritannien, Kanada, USA). Diese Kurse sollten fachspezifisch ausgerichtet sein (Weiterbildungsassistenten jeweils einer Fachrichtung) und neben den klassischen Inhalten besonders den Beitrag von EbM zur alltäglichen Problemlösung hervorheben: durch praktische Übungen wie auch das Lernen am kollegialen Vorbild. Die Weiterbildung für Allgemeinmedizin hat in den letzten zwölf Jahren gezeigt,

dass ein Kursangebot als Ergänzung zur praktischen Weiterbildung sinnvoll ist. Die Weiter- und Fortbildungsakademien sollten, allerdings zunächst auf freiwilliger Basis, dieses Prinzip für EbM-Inhalte auf sämtliche Fachrichtungen ausdehnen.

Qualitativ und quantitativ sind Angebote für die zu Beginn genannten Meinungsbildner zu erweitern. Hier sind durchaus Initiativen vorhanden, z.B. vonseiten der Kassenärztlichen Vereinigungen, die allerdings intensiviert und verstetigt werden müssen.

Aber es kommt nicht nur auf die „Software“ an, also die edukativen Angebote und Bemühungen von Kliniken und Praxen, die Logik der EbM zu integrieren. Vielmehr muss auch die „Hardware“ weiter entwickelt werden, d.h. die Datenbanken bzw. Medien, auf die ein Arzt zurückgreifen kann, der eine fundierte Antwort auf seine Frage sucht. Hier ist noch viel zu tun, um Validität, Zugänglichkeit, Benutzerfreundlichkeit und Umfang zu verbessern. Ärzte werden hier Enthusiasmus, Zeit und Geld (Abonnementsgebühren!) investieren müssen.

Literatur

- [1] Sackett DL et al. (1997) Evidence-Based medicine. How to practice and teach EbM. Churchill-Livingstone, New York
- [2] Fritsche L et al., Do short courses in evidence based medicine improve knowledge and skills? Validation of Berlin questionnaire and before and after study of courses in evidence based medicine. *BMJ*. (2002), 325, 1338–1341
- [3] Bateman H, Flinn T, Exploring aspects of evidence-based medicine: case report of a practice-based approach to learning. *Educ Gen Pract*. (1999), 10, 119–123
- [4] Coomarasamy A, Khan KS, What is the evidence that postgraduate teaching in evidence based medicine changes anything? A systematic review. *BMJ*. (2004), 329, 1017
- [5] Brassey J et al., Just in time information for clinicians: a questionnaire evaluation of the ATTRACT project. *BMJ*. (2001), 322, 529–530

41 Schulung von medizinischen Laien

Bettina Berger, Nicole Skoetz, Olaf Weingart

41.1 Patientenschulung in EbM – ein reales Szenario

Reiner F. erhielt für seine starken Stimmungsschwankungen vor 10 Jahren die Diagnose „Bipolare Störung“. Als Architekt war er eigenständiges Arbeiten gewohnt. Die Erkrankung raubte ihm seine berufliche Existenz. Es dauerte Jahre, bis er nach diversen Klinikaufenthalten die für ihn angemessene Behandlungsform finden konnte. Letztendlich verdankt er den Mut für die Anforderung einer ihm angemessenen Therapie der Selbsthilfegruppe, in der er seit Jahren tätig ist. Nun möchte er gerne die Lehren, die er aus seiner Erkrankung ziehen konnte, allen in ähnlicher Weise Betroffenen zugänglich machen. Er ist überzeugt, dass Selbsthilfe ein ausgesprochen wichtiger Bestandteil jedes Behandlungskonzeptes sein sollte. Deshalb möchte er sich dafür einsetzen, dass die Empfehlung zur Mitarbeit in einer Selbsthilfegruppe in die Leitlinien (s. Kap. 35) aufgenommen wird. Er sucht nach Möglichkeiten der Weiterbildung, um sich aus der Betroffenenperspektive in die Auseinandersetzung um eine patientenorientierte Versorgung einbringen zu können. Welche Fortbildungsangebote könnten ihm dabei behilflich sein, sich das notwendige Wissen für eine aktive Mitarbeit in einer Leitliniengruppe anzueignen?

41.2 EbM-Training für Laien

Training für Laien richtet sich an die Nutzer des Gesundheitswesens im Sinne der Definition des **Sachverständigenrates für Konzertierte Aktion im Gesundheitswesen**. Als „Nutzer“ wird hier jede Person bezeichnet,

die Zugang zum System der gesundheitlichen Versorgung hat, unabhängig davon, ob dieser Zugang aktuell genutzt wird oder nur fakultativ besteht. Ohne die aktive Mitwirkung der Leistungsadressaten, also der Laien bzw. Patienten, kommt das moderne Gesundheitssystem nicht mehr aus – sagt die Selbsthilfeforschung. Sie hat erkannt, dass das Potenzial der Laien eine soziale, politische und ökonomische Ressource für das System der Gesundheitssicherung darstellt [1]. Dabei ist zu berücksichtigen, dass Menschen verschiedene Rollen innehaben können – als Bürger, Versicherte, Patienten – und hier wieder als Koproduzenten ihrer eigenen Gesundheit, als Bewertende des Gesundheitswesens, als Kunden auf dem Gesundheitsmarkt, als Partner im medizinischen Entscheidungsprozess oder als zu versorgende, passive Patienten. Ihre Expertise als Betroffene können und wollen diese Personen unterschiedlich intensiv in die Entscheidungsprozesse einbringen. Den Laienstatus behalten sie gegenüber dem Leistungsanbieter in Bezug auf die Systemzugehörigkeit und den damit verbundenen Wissensvorsprung und die Handlungsgewalten.

41.3 Warum EbM-Training für Laien?

Zunehmend können und wollen sich Patienten mit dem Status „Betroffener“ allein nicht mehr zufrieden geben. Wie andere Laienbewegungen auch – so die Umwelt-, Frauen- oder Friedensbewegung – wollen sie mit über das System der Gesundheitsversorgung

nachdenken und entscheiden. Die WHO hat bereits 1994 Bürgerbeteiligung in ihre „Declaration on the promotion of Patients' Rights in Europe“ aufgenommen und dort ausgeführt [2]: „Patients have a collective right to some form of representation at each level of the health care system in matters pertaining to the planning and evaluation of service, including the range, quality and functioning of the care provided.“

In Deutschland fordert das Sachverständigen-gutachten von 2001 erstmals offiziell eine verstärkte Patientenorientierung und -einbeziehung in das Gesundheitswesen [3]. Für die Wahrnehmung dieser kollektiven Patientenrechte sind Training und Schulung notwendig – so fordert es nicht nur Christoph Kranich in seinem Artikel „Welche Kompetenzen brauchen Patienten?“ mit den Worten: „Wenn die Medizin zunehmend wissenschaftsbasiert sein will, müssen die Betroffenenvertreter wissen, worum es geht, um kompetent in den Gremien des Gesundheitswesens mitarbeiten zu können.“ [4]

Das Konzept der EbM hat den Anspruch, Entscheidungen auf der Basis

- ▲ der individuellen (klinischen) Erfahrung des Arztes,
- ▲ der Bedürfnisse des Patienten und
- ▲ der bestmöglichen verfügbaren, externen Evidenz aus systematischer Forschung zu treffen.

Um dieses Konzept auf Laien anzuwenden, können auch sie in die Lage versetzt werden,

- ▲ zu ihren individuellen Bedürfnissen suchtaugliche Fragen zu stellen,
- ▲ nach Wissen (externer Evidenz) zu suchen,
- ▲ diese Informationen zu bewerten und einzuordnen sowie
- ▲ die Erkenntnisse in den Entscheidungsprozess einfließen zu lassen.

Auch wenn die zunehmende Verfügbarkeit von „Evidenzbasierten Patienteninformationen“ (s. Kap. 18) in vielen Fällen des Alltags jetzt schon hilfreich sein dürfte, erfordert

eine erfolgreiche Umsetzung von EbM langfristig die Kooperation und Akzeptanz der Patienten. Voraussetzung hierfür ist ein Grundlagenverständnis der Bürger zur Bedeutung der Wissenschaftsbasierung von Informationen zu Gesundheits- und Krankheitsfragen. Aus dieser Erkenntnis heraus wurden in den letzten Jahren in Deutschland von verschiedenen Organisationen insbesondere für Patientenvertreter und andere Multiplikatoren EbM-Schulungsangebote entwickelt.

Auf der Meso- und Makroebene der Mitarbeit von Patientenvertretern in Strukturen des Gesundheitswesens geht es darum, eine patientenorientierte Gesundheitsversorgung zu etablieren und in den Gremien fachkompetent mitarbeiten zu können. Konkret handelt es sich um die Einbeziehung von Patienten in Strukturen wie Leitlinienkommissionen, Ethikkommissionen, Gesundheitskonferenzen, den Gemeinsamen Bundesausschuss etc. Seit der Gesundheitsreform 2004 partizipieren über 1000 Patientenvertreter auf Bundes- und Landesebene in den verschiedenen Gremien. Hier sind Grundlagenkenntnisse des Gesundheitswesens und der EbM notwendig.

41.3.1 Internationale Erfahrungen

Internationaler Vorläufer für das Kompetenztraining der Fachwissenschaft Gesundheit der Universität Hamburg ist das LEAD-Projekt der Brustkrebsorganisation der USA. LEAD steht für Leading Education and Advocacy Development und bietet ein einwöchiges Trainingsprogramm an. Hier werden seit 10 Jahren von Brustkrebs betroffene Frauen in wissenschaftlichen Kompetenzen geschult, um aktiv an der Gestaltung der Versorgung von Brustkrebspatientinnen und der Forschung in diesem Gebiet teilzunehmen¹.

¹ Der Einfachheit halber wird im Falle der Brustkrebspatientinnen die weibliche Form gewählt. Betroffene Männer sind natürlich mitgemeint.

Ca. 1000 Absolventinnen weist dieses Projekt bereits auf, die in jährlichen Absolvententagungen Möglichkeiten zur Weiterbildung und zum Austausch finden. 2004 fand das erste internationale Trainingsprogramm in Madrid mit Teilnehmerinnen aus 24 Ländern weltweit statt. Absolventinnen der LEAD-Programme bringen sich erfolgreich auf den verschiedenen Ebenen des Gesundheitswesens ein und sind spürbar beteiligt an der Entwicklung von Studiendesigns zur Brustkrebstherapie [5].

In Großbritannien wurde das Critical-appraisal-skills-Programm (CASP) ursprünglich zur Schulung von Laien und Verbrauchern entwickelt [6]. Das Programm arbeitet mit dem Ziel, wissenschaftliche Literatur kritisch zu bewerten und auf die jeweiligen Fragen des Gesundheitswesens anzuwenden. Auch das Scottish Intercollegiate Guidelines Network (SIGN) bietet eintägige Workshops für Interessenten zur Mitarbeit in Leitliniengruppen an.

41.4 Projekte zur EbM-Laienschulung in Deutschland – mehr als ein theoretisches Modell

Dem Bedarf entsprechend und mithilfe finanzieller Förderung entwickelten und evaluierten drei Arbeitsgruppen in Deutschland EbM-Schulungsmaßnahmen für Laien. Parallel zu und ausgehend von diesen drei Modellprojekten entstanden weitere Angebote der EbM-Laienfortbildung (s. Tab. 41.1).

1. Training in wissenschaftlicher Kompetenz von Patienten- und Verbrauchervertretern (Fachwissenschaft Gesundheit, Universität Hamburg)

Das Kompetenztraining verfolgte drei Zielsetzungen:

- ▲ die Qualifizierung von Patientenberatern in Bezug auf die Entstehung, Bewertung und Erstellung bzw. Einforderung evidenzbasierter Patienteninformationen,

Tab. 41.1: Bisherige EbM-Schulungsangebote für medizinische Laien in Deutschland

Name	Zielgruppe	Inhalte/Themenschwerpunkte	Umfang	Organisation
„Curriculum Patientenberatung“ – Suchen, Bewerten und Vermitteln von Patienteninformationen –	Personen unterschiedlicher Berufsgruppen, die in der Beratung zu Gesundheitsfragen tätig sind.	<ul style="list-style-type: none"> • Ermitteln des Informationsbedarfs des Klienten • Suche nach interner und externer Evidenz • Kritische Bewertung von Informationen • Informationsgestützte Kommunikation des Beraters mit dem Klienten • Rechtsgrundlagen der Patientenberatung 	ca. 20 h (Teilmodule möglich)	<p>Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Berlin (www.patienten-information.de)</p> <p>In Kooperation mit der Abteilung Epidemiologie, Sozialmedizin und Gesundheitssystemforschung der MH Hannover und dem deutschen Cochrane Zentrum</p>
Patientenworkshop Schulungsprogramm für Patienten zur Mitarbeit bei der CHMG.	Patienten, Angehörige, Vertreter von Selbsthilfegruppen und Patientenorganisationen (vorzugsweise Onkologie)	<ul style="list-style-type: none"> • Einführung in EbM • Statistik • Bewertung von Patienteninformationen, Studien, Protokollen und Systematischen Reviews 	ca. 22 h	<p>Cochrane Haematological Malignancies Group (CHMG)</p> <p>Klinikum der Universität zu Köln (www.chmg.de)</p>

Tab. 41.1: Fortsetzung

Name	Zielgruppe	Inhalte/Themenschwerpunkte	Umfang	Organisation
Kompetenztraining für Patienten- und VerbrauchervertreterInnen	Mitglieder und Mitarbeiter aus Selbsthilfegruppen, Patienteninitiativen, und anderen Einrichtungen der Patientenunterstützung und -beratung, sowie Patientenvertreter in Organisationen des Gesundheitswesens	<ul style="list-style-type: none"> • EbM für Patienten- und Verbrauchereretreter • Statistik • Formulierung einer suchtauglichen Fragestellung • Einführung in medizinischen Datenbanken (PubMed, Cochrane) • Kritische Bewertung: (Beobachtungsstudie, randomisiert-kontrollierte Studie, diagnostische Studie, systematische Übersichtsarbeiten) • Evidenzbasierte Patienteninformation Patientenbeteiligung am Beispiel der Ethikkommission 	ca. 40 h	Fachwissenschaft Gesundheit – Universität Hamburg (www.gesundheit.uni-hamburg.de)
ebm@school Modellprojekt zur Entwicklung und Erprobung eines Curriculums zur Kompetenzsteigerung der critical health literacy	Schülerinnen und Schüler der 11. Klassen an Allgemeinbildenden Schulen.	<ul style="list-style-type: none"> • http://gesundheit.chemie.uni-hamburg.de/cgi-bin/newsite/index.php?page=page_331 	ca. 22 h	Fachwissenschaft Gesundheit – Universität Hamburg
MS-Training	Patientinnen mit schubförmiger MS	<ul style="list-style-type: none"> • Lesen, Interpretieren und Aufarbeiten aktueller Studienergebnisse klinischer Studien zur Interferontherapie 	ca. 4 h	Fachwissenschaft Gesundheit, UKE Hamburg
Qualifizierungsprogramme für Journalisten	Journalisten und Medienvertreter	<ul style="list-style-type: none"> • Einführung in EbM • Nutzung von EbM-Quellen (z.B. Cochrane Library) 	„EbM-Teil“ unterschiedlich	Diverse Stiftungen (u.a.: Bertelsmann, Heureka) in Zusammenarbeit mit dem Deutschen Cochrane Zentrum

- ▲ die Qualifizierung der Patientenvertreter zur Teilnahme an medizinischen Entscheidungen auf der Meso-Ebene (Patient ↔ Institution),
- ▲ Ausbildung von Multiplikatoren zur Etablierung eigener Schulungsprogramme.

Daraus leiten sich die angesprochenen Zielgruppen des Trainings ab: Patientenvertreter aus Selbsthilfegruppen (ehrenamtlich und hauptamtlich), Patientenberater und Multiplikatoren. Ausgeschlossen waren Mediziner und Interessenten, die nicht am gesamten Kurs teilnehmen konnten bzw. solche, die den Kurs nur für die persönliche Weiterbildung nutzen wollten.

Im Rahmen der Förderung über § 65b SGB V konnten 14 Schulungen über fünf Tage bzw. zwei mal zweieinhalb Tage mit jeweils 40 Stunden durchgeführt werden. 161 Teilnehmer der genannten Zielgruppen wurden erreicht. Die Schulungen wurden in den Räumen der Universität Hamburg durchgeführt. Zur Vorbereitung auf den Kurs wurde ein Reader mit den zu bearbeitenden Studien verschickt.

2. Schulungsprogramm für Patienten zur Mitarbeit bei der Cochrane Haematological Malignancies Group (CHMG)

Das von der Deutschen Krebshilfe geförderte Schulungsprogramm der CHMG richtete sich primär an Patientenvertreter, die sich aktiv in die wissenschaftliche Aktivität der CHMG einbringen wollten. Es galt, ein Patientennetzwerk aufzubauen, das einerseits die von der CHMG erstellten Protokolle und Reviews im Hinblick auf Patienten-Relevanz kommentiert, andererseits auch auf Basis dieser Review-Ergebnisse Informationen für Patienten erstellt.

Das Konzept der Patientenworkshops der CHMG basierte auf den Erfahrungen verschiedener nationaler und internationaler Arbeitsgruppen, die sowohl Ärzte als auch

medizinische Laien als Zielgruppe ansprechen. Die begleitende Evaluation der durchgeführten Workshops zeigte, dass ein solches Konzept zur Schulung in EbM auch bei Patienten erfolgreich umgesetzt werden kann. Von 21 in zwei Grund- und einem Aufbaukurs geschulten Patienten beteiligten sich acht in der Folgezeit aktiv an der Arbeit der CHMG.

3. Curriculum Patientenberatung des Ärztlichen Zentrums für Qualität in der Medizin (ÄZQ)

Einen anderen Ansatz verfolgten die im Rahmen eines Modellprojektes beim ÄZQ durchgeführten Kurse für Patientenberater (www.patienten-information.de). Dieses ebenfalls nach § 65b SGB V geförderte Modellprojekt sollte die Qualität der Patientenberatung durch Nutzung qualitätsgeprüfter Informationen verbessern. Das begleitend erstellte Curriculum zur Ausbildung von Patientenberatern enthält u.a. Module zum Umgang mit der Internettechnologie, dem Patienteninformationssystem und dem Qualitätsbewertungsinstrument für Patienteninformationen DISCERN sowie spezielle Aspekte der Kommunikation der Evidenz [7].

Weiterführende Literatur zu den Modellprojekten [7; 11; 13; 15; 16].

41.4.1 Curriculare Entwicklung eines EbM-Trainings für Laien am Beispiel des Trainings in wissenschaftlicher Kompetenz für Patienten- und Verbrauchervertreter der Universität Hamburg

Die Fachwissenschaft Gesundheit der Universität Hamburg beschäftigt sich seit Jahren mit der Entwicklung, Durchführung und Evaluation von EbM-Kursen für Pflegende, Diabetesberater, Studierende der Fachwissenschaft Gesundheit und Ärzte. Auf diesen, wie auch den internationalen Erfahrungen aus

dem CASP und dem LEAD-Projekt konnte das Curriculum für das „Kompetenztraining für Patienten- und VerbrauchervertreterInnen“ aufbauen. Zahlreiche einzelne Bausteine und didaktische Hilfsmittel zur Vermittlung der Grundlagen der EbM lagen bereits vor. So veranschaulicht z.B. ein Glas mit farbigen Perlen die zufällige Verteilung von Interventionseffekten und wird außerdem zur Veranschaulichung der Unterscheidung von absoluter und relativer Risiko-Reduktion verwendet. Für andere Unterrichtsmodule wurden grundlegende theoretische Bausteine im Verlauf des Projektes entwickelt, z.B. Kriterien zur Bewertung von evidenzbasierten Patientinformationen [8]. Auf dieser Grundlage konnten wiederum wissenschaftsbasierte Bausteine für den Unterricht entwickelt werden, die wie die Vierfeldertafel im Rahmen einer Studie [8] im Vergleich zum Entscheidungsbaum auf ihr Verständnis hin überprüft wurden. Einzelne Module sind darüber hinaus unabhängig von einer Kursteilnahme online zu nutzen (www.gesundheit.uni-hamburg.de).

Die Entwicklerinnen des Curriculums nahmen an internationalen EbM-Kursen von CASP in Oxford und dem LEAD-Projekt teil, um deren Erfahrungen in den Kurs zu integrieren. Berücksichtigt wurden insbesondere Prinzipien der Erwachsenenpädagogik und des problemorientierten Lernens. So wurden alle Kursteilnehmer nach ihren individuellen Lernzielen befragt, um diese im Laufe der Kurswochen zu berücksichtigen und ihnen die Bewertung des Kurses anhand der individuellen Lernziele zu ermöglichen. Vorhandene Vorerfahrungen der Teilnehmer konnten bei der Kleingruppenarbeit angemessen zur Geltung kommen. Während der Unterrichtseinheiten wurde auf die Verwendung unterschiedlicher Methoden geachtet, z.B. Präsentationen, Unterrichtsgespräche, Kleingruppenarbeit mit vorbereiteten Arbeitsaufträgen, Anwendungsübungen und eigenständige Rechercheübungen an Computerarbeitsplätzen.

Am Beispiel der Hormontherapie in den Wechseljahren wurde zuerst der öffentliche Diskurs in den Medien dargestellt, dann die Nurses' Health Study als Beobachtungsstudie gelesen, um darauf aufbauend die Überlegenheit der randomisiert-kontrollierten Studie darstellen zu können. In Kleingruppen wurden die Studien diskutiert und die ungeklärten Fragen gesammelt. In der Einheit „Formulierung einer suchtauglichen Fragestellung“ konnten mit Hilfe des PIKE-Schemas eigene Fragen entwickelt und in PubMed (www.pubmed.gov) nach relevanter Literatur recherchiert werden. Ein weiteres Modul veranschaulichte die Kommunikation von Studienergebnissen mit graphischen Hilfsmitteln auf der Grundlage der errechneten Zahlen.

In einem abschließenden Test wurden diese Elemente wiederholt und überprüft: „Formulierung einer suchtauglichen Fragestellung“, „Lesen und Verstehen einer Studie“, „Berechnen wesentlicher Ergebnisparameter“ und „Kommunikation der Ergebnisse“. Dieser Test diente somit sowohl der Evaluation des Kurses als auch der Selbstevaluation der erworbenen Kompetenzen.

41.5 Ergebnisse

Trotz teilweise unterschiedlicher Zielsetzungen dieser evaluierten Modellprojekte wird durchgängig berichtet, dass die Teilnehmer die Partizipation als positiv und für ihre Arbeit als hilfreich erachteten.

Der in Hamburg durchgeführte Kompetenztest zeigte, dass der Durchschnitt der erreichten Ergebnisse vergleichbar ist mit dem der Studierenden der Fachwissenschaft Gesundheit nach einem Semester „Wissenschaftsliteraturseminar“ mit entsprechendem Inhalt [9]. Damit konnte belegt werden, dass medizinische Laien ausreichend motiviert sind, um sich die Grundlagen der EbM anzueignen. Die qualitativen Feedbackmel-

dungen der Teilnehmer nach Kursteilnahme ergaben neue Erkenntnisse in folgenden Dimensionen: Ein selbstbewussteres Rollenverständnis, eine Ernüchterung gegenüber der Wissenschaftlichkeit der Medizin, die Erkenntnis, dass EbM als Methodik nutzbar und verfügbar ist, dass der Umgang mit Zahlen die eigene Kompetenz steigert, dass EbM als Handwerkszeug zur Studienbewertung auch von Laien erlernbar ist, erhöhte Recherchekompetenzen und neue Konzepte der Patientenberatung und -information.

Die Lernbarrieren sind mit den Erfahrungen von EbM-Kursen für Ärzte [10] vergleichbar:

- ▲ teilweise englischsprachiges Material nutzen zu müssen,
- ▲ unterschiedliche Kenntnisse und Erfahrungen im Umgang mit weiteren Hilfsmitteln, insbesondere dem Gebrauch von Computern und Internet,
- ▲ unterschiedliches Wissen über die behandelte Fragestellung, wobei bei Patienten die Informationsbewertung sowohl von den Unterschieden im Inhaltswissen (Fachwissen) als auch durch die mitunter starke persönliche Betroffenheit beeinflusst wird.

Auch wenn den EbM-Grundkursen für Ärzte vergleichbare Inhalte angeboten wurden, erwies sich eine gemeinsame Schulung von Patienten-(Vertretern) und (klinisch tätigen) Ärzten zur Vermittlung von EbM-Grundlagen meist nicht als sinnvoll, da sich die differierenden Handlungskonzepte bei den zu entwickelnden Fragestellungen und den aus den Ergebnissen zu ziehenden Schlüssen als Störfaktor erwiesen [11]. Bei allen Programmen erwies sich die Orientierung an den Fragestellungen der Teilnehmer als hilfreich für die eigenständige Umsetzung.

Die erfolgreiche Umsetzung von Kompetenzen wurde entweder von strukturellen Barrieren limitiert oder erst nach langer Zeit sichtbar. Beispiele für Umsetzungen der im

Kompetenztraining für Patienten- und VerbrauchervertreterInnen erworbenen Fähigkeiten:

- ▲ MS-Journalclub (Journal-Club von Patienten mit Multipler Sklerosis),
- ▲ Mitarbeit in Ethikkommission,
- ▲ Mitarbeit als Patientenvertreter in KTQ-Arbeitsgruppen (Kooperation für Transparenz und Qualität im Krankenhaus) und Gemeinsamen Bundesausschuss (G-BA),
- ▲ Betroffenenvertretung in Studie zur Angehörigenbetreuung von Alzheimer-Patienten,
- ▲ Mitarbeit bei Leitlinienerstellung Bipolare Störungen,
- ▲ Integration von EBM-Inhalten in das Curriculum der Familienberaterausbildung der WHO,
- ▲ Workshops mit ACHSE e.V. zum Thema seltene Erkrankungen und EbM,
- ▲ Entwicklung von Kompetenztrainings für Brustkrebsaktivistinnen (KOMBRA) und Österreich.

Diese Erfolge bestätigten sich im Zuge der Projektevaluation des „Curriculum Patientenberatung“ des ÄZQ. Hier wurden in einem kontrollierten Design mit Interventions- und Wartekontrollgruppe 38 Patientenberater aus Gesundheitsämtern (n=11), Selbsthilfeeinrichtungen (n=13) und kommunal oder über Verbände finanzierten Beratungsstellen (n=14) geschult. Nach der Schulung waren in der Nutzerbefragung der Beratungsstellen Verbesserungen bei der überwiegenden Zahl der Items erkennbar. Bei der differenzierten Auswertung nach der Art der Beratung zeigten sich die deutlichsten Effekte für Einrichtungen der Selbsthilfe mit signifikanten Verbesserungen in fast allen Items gegenüber nahezu unveränderten Bewertungen bei den übrigen Einrichtungen [12]. Die Nachhaltigkeit der Umsetzung der erworbenen Fähigkeiten und Fertigkeiten könnte durch Aufbau- und Vertiefungsmaßnahmen

dule sowie durch die Ergänzung mit Online-modulen erhöht werden. Auch wenn in vielen Bereichen insbesondere die Selbsthilfevertreter die erworbenen Kompetenzen umsetzen konnten und oft auch lokal als Multiplikatoren fungierten, ist das Angebot der langfristigen Begleitung und Unterstützung der EbM-Arbeit von Laien von Bedeutung.

So erwies sich bei der CHMG-Arbeitsgruppe als effektivste Maßnahme zur Sicherstellung der Begutachtung von Reviews die Durchführung eines Seminars mit einer Fokusgruppe, die diese gemeinsam bewerkstelligte. Die hier mitarbeitenden Patienten wurden Teil des Teams, das die Begutachtung der von den Autoren eingereichten Reviews und Protokolle durchführte [13].

Auflösung des Szenarios

Reiner F. hat den Kurs in der Fachwissenschaft Gesundheit absolviert. Die Grundlagenkenntnisse, die er dort erwerben konnte, reichten bei weitem nicht aus. Immer wieder kam er zu Beratungen in die Uni. Im Laufe von drei Jahren entwickelten sich Ideen und Konzepte. Derzeit arbeitet er an einer Literaturrecherche zur Evidenz von Selbsthilfe bei Bipolaren Störungen. Gerne würde er weitere Kurse speziell zum Thema „Mitarbeit in Leitliniengruppen“ besuchen.

41.6 Ziele in der Zukunft

Derzeit besteht für keines der o.g. Schulungsprogramme für Laien eine weitere finanzielle Förderung, sodass eine wissenschaftlich fundierte Weiterentwicklung dieser Programme nicht erfolgt. Lediglich das MS-Schulungsprogramm wird derzeit umgesetzt, während in anderen Bereichen insbesondere die Erfahrungen der an den Projekten beteiligten Referenten im Rahmen von Weiterentwicklungen von Schulungen für Laien genutzt werden. So unterstützt z.B. die Fachwissenschaft Gesundheit der Uni Hamburg zurzeit

das Kombra-Projekt (Kompetenztraining Brustkrebsaktivistinnen).

Die hier beschriebenen Projekte stellen nur einen Anfang dar. Langfristig geht es darum, sensibel für die Lücken zwischen EbM und Patientenorientierung zu werden und diese beiden Ansätze miteinander in Verbindung zu bringen [14]. Es kann also nicht nur darum gehen, Patienten zu vermitteln, was unter EbM zu verstehen ist, sondern es geht darum, mit deren Unterstützung an einem patientenorientierten Gesundheitswesen zu arbeiten. Wie in anderen Ländern sollte jede Organisation, in der sich Patienten an Entscheidungen zu Gesundheitsfragen beteiligen, gewährleisten, dass diese Vertreter eine ausreichende inhaltlich und methodische Kompetenz erwerben können. Daher sollte es ein Ziel der Zukunft sein, aufbauend auf den Erfahrungen der Schulungsprogramme nachhaltige Aus- und Weiterbildungsstrukturen für Patienten- und Verbrauchervertreter zu entwickeln und zu implementieren, deren Aufgabe es u.a. ist:

- ▲ Entwicklungsprozesse klinischer Studien zu begleiten,
- ▲ in Leitliniengruppen mitzuarbeiten und
- ▲ evidenzbasierte Patienteninformationen zu erstellen.

Diese sollten die Fragestellungen der Patienten- und Verbrauchervertreter aufgreifen und möglichst frei von Interessenkonflikten sein. Darüber hinaus erfordert eine evidenzbasierte Lösung von Gesundheitsfragen von und mit dem Patienten auch eine konzeptionelle Einbindung in die Gesundheitsaufklärung und Gesundheitsbildung von Patienten und Bürgern. Neben Bereitstellung entsprechender Medien würde dies in Analogie zu Train-the-Trainer-Seminaren Lern- und Lehrangebote für Multiplikatoren im Bereich der Selbsthilfe wie auch für Lehrer an Allgemeinbildenden Schulen erfordern.

Literatur

- [1] Trojan A (2003) Der Patient im Versorgungsgeschehen: Laienpotential und Gesundheitsselbsthilfe. In: Schwartz FW et al., Public Health. Gesundheit und Gesundheitswesen, 321–333. Urban und Fischer, München
- [2] World Health Organization (WHO). A declaration on the promotion of patients' rights in Europe. European consultation on the right of patients. Amsterdam 1994. http://www.who.int/genomics/public/eu_declaration1994.pdf (14.12.06)
- [3] Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen (2001) Bedarfsgerechtigkeit und Wirtschaftlichkeit. Band I: Zielbildung, Prävention, Nutzerorientierung und Partizipation. Band II: Qualitätsentwicklung in Medizin und Pflege. Gutachten 2000/2001. Kurzfassung., Bonn
- [4] Kranich C, Patientenkompetenz. Was müssen Patienten wissen und können? Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. (2004), 47, 950–956
- [5] Dickersin K et al., Development and implementation of a science training course for breast cancer activists: Project LEAD (leadership, education and advocacy development). Health Expect. (2001), 4, 213–220
- [6] Milne R, Oliver S, Evidence-based consumer health information: developing teaching in critical appraisal skills. Int J Qual Health Care. (1996), 8, 439–445
- [7] Dierks ML, Huth A, Lang B, Maienborn A, Martin S, Moroff J, Sängler S, Weingart O, Glowienka-Wiedenroth F, Lerch M, Curriculum Patientenberatung. Informationen recherchieren, bewerten, vermitteln. 2nd ed. http://www.patienten-information.de/content/projekte/kontinuierlich/download/curriculum_ebm_patinfo_akt.pdf (12.12.06)
- [8] Steckelberg A et al., Kriterien für evidenzbasierte Patienteninformatoren. Z Arztl Fortbild Qualitätssich. (2005), 99, 343–351
- [9] Berger B, Kasper J, Mühlhauser I, Training in wissenschaftlicher Kompetenz von Patienten- und Verbrauchervertretern. Evaluation eines Messinstrumentes, (Abstract). Z Arztl Fortbild Qualitätssich. (2004), 625
- [10] Meyer T, Ströbel A, Raspe H, Evidenzbasierte Medizin aus Sicht niedergelassener Ärztinnen und Ärzte: ein repräsentativer Survey zu Akzeptanz und Fortbildungsinteressen. Z Arztl Fortbild Qualitätssich. (2004), 98, 293–300
- [11] Weingart O et al., Patientenbeteiligung in der Cochrane Collaboration-Barrieren, Erfahrungen und Konzepte aus deutscher Sicht. Z Arztl Fortbild Qualitätssich. (2005), 99, 367–371
- [12] Martin S, Sängler S, Dierks ML, Effekte des Schulungsprogramms „Curriculum Patientenberatung“ auf die Wahrnehmung der Beratungsqualität durch Klientinnen und Klienten der teilnehmenden Beratungsstellen (Abstract). Gesundheitswesen. (2003), 65, A1–A88
- [13] Skoetz N, Weingart O, Engert A, A consumer network for haematological malignancies. Health Expect. (2005), 8, 86–90
- [14] Bensing J, Bridging the gap. The separate worlds of evidence-based medicine and patient-centered medicine. Patient Educ Couns. (2000), 39, 17–25
- [15] Weingart O, Skoetz N, Wenn Patienten lernen Ergebnisse medizinischer Forschung zu bewerten. Qualitas. (2004), 16–18
- [16] Sängler S et al., Förderung des ÄZQ als Einrichtung der unabhängigen Verbraucher- und Patientenberatung. Sachstandsbericht. Z Arztl Fortbild Qualitätssich. (2005), 99, 406–409

VI Deutsche EbM-Organisationen

42	Das Deutsche Netzwerk Evidenzbasierte Medizin (DNEbM)	439
	<i>Jürgen Windeler</i>	
43	Cochrane Collaboration: Deutsches Cochrane Zentrum/ Deutsche Cochrane-Review-Gruppen	443
	<i>Antje Timmer, Gerd Antes</i>	
44	Das ÄZQ – Kompetenzzentrum der ärztlichen Selbstverwaltung für Qualität und Evidenz in der Medizin	457
	<i>Günter Ollenschläger, Christian Thomeczek, Sylvia Säger, Monika Lelgemann</i>	
45	Das Netzwerk der Koordinierungszentren für Klinische Studien	467
	<i>Christoph M. Seiler, Insa Bruns, Stephanie Wolff, Walter Lehmacher, Markus Löffler, Herbert Maier-Lenz</i>	
46	DIMDI – Deutsches Institut für Medizinische Dokumentation und Information ...	475
	<i>Alric Rüther, Britta Göhlen</i>	

42 Das Deutsche Netzwerk Evidenzbasierte Medizin (DNEbM)

Jürgen Windeler

Nachdem Ende der 80er- und Anfang der 90er-Jahre die Idee der Evidenzbasierten Medizin aus den angelsächsischen Ländern langsam nach Deutschland diffundierte und Mitte der 90er-Jahre die ersten Kurse zur Einführung in Evidenzbasierte Medizin veranstaltet wurden, gründete sich im März 1998 in Berlin eine Arbeitsgemeinschaft „Deutsches Netzwerk Evidenzbasierte Medizin“ [1]. Sie bestand aus einer Reihe von Fachleuten und Interessierten aus verschiedenen Professionen des Gesundheitswesens. Ihr Ziel war es, den Gedanken der Evidenzbasierten Medizin in Deutschland zu verbreiten und strukturiert weiterzuentwickeln. Aus dieser Arbeitsgemeinschaft heraus wurde am 6. Oktober 2000 der gleichnamige Verein „Deutsches Netzwerk Evidenzbasierte Medizin e.V.“ (DNEbM) gegründet. Die Gründung ging einher mit weiteren Meilensteinen im Jahr 2000, der Einführung des Begriffes „Evidenzbasierte Leitlinien“ in das SGB V (in Zusammenhang mit den so genannten Chroniker-Programmen), dem Erscheinen der ersten Auflage des Lehrbuches „Evidenzbasierte Medizin in Klinik und Praxis“ [2] sowie dem zweiten Symposium Evidenzbasierte Medizin in Berlin, in dessen Zusammenhang die Gründung des Vereins erfolgte. Seit Juni 2000 hat die Arbeitsgemeinschaft/der Verein als sein offizielles Publikationsorgan die Zeitschrift für ärztliche Fortbildung und Qualität im Gesundheitswesen (ZaeFQ).

Ausgehend von den wenigen Gründungsmitgliedern hat sich das DNEbM in den letzten 6 Jahren zu einer veritablen wissenschaftlichen Fachgesellschaft mit ca. 700 Mitgliedern (Stand: Ende 2006) entwickelt.

Seit November 2004 ist der Verein Mitgliedsgesellschaft der Arbeitsgemeinschaft der Wissenschaftlich Medizinischen Fachgesellschaften (AWMF).

42.1 Zielsetzungen

Die übergreifende Zielsetzung der Gründung des Netzwerkes war eine Plattform zum Austausch und zur Information für alle diejenigen, die sich mit Evidenzbasierter Medizin beschäftigen oder sich dafür interessieren. Wenn man sich vergegenwärtigt, dass es sich bei Evidenzbasierter Medizin weniger um eine eigenständige wissenschaftliche Disziplin als vielmehr um ein übergreifendes wissenschaftliches Konzept unter Einbindung von Aspekten verschiedener Fachrichtungen handelt, so wird verständlich, dass es spezieller Anstrengungen bedurfte und bedarf, die an EbM Interessierten, die sich auf verschiedene Fachdisziplinen und verschiedene Fachgesellschaften und Institutionen verteilen, zusammenzubringen, um so gemeinsamen Interessen nachzugehen. Die Zielsetzungen des Netzwerkes lassen sich daher vor allem in folgenden Punkten zusammenfassen:

- ▲ die Integration der Mitglieder des Netzwerkes und der sie tragenden wissenschaftlichen Disziplinen und Arbeitsbereiche,
- ▲ der Aufbau einer offenen Informations- und Kommunikationsplattform,
- ▲ die Entwicklung von Aus-, Weiter- und Fortbildungscurricula und -modellen,
- ▲ die Abstimmung laufender EbM-bezogener Forschung, Ausbildung und Praxis,

- ▲ die Durchführung von Evaluations- und Forschungsprojekten,
- ▲ die Weiterentwicklung von Theorie, Methoden und Ethik der EbM.

Zu weiteren Zielsetzungen kann man die internationale Kooperation sowie die Einflussnahme auf Strukturen des Gesundheitssystems ebenso zählen wie die Öffentlichkeitsarbeit, ganz speziell mit der Zielrichtung der wissenschaftlichen Ausrichtung der Redaktionen der Printmedien.

42.2 Struktur

Das DNEbM, das seit 2000 über eine eigene ansprechende und informative Website (<http://www.dnebm.de>) verfügt, ist in mehrere Fachbereiche unterteilt, die die Interessen und Arbeitsschwerpunkte der Mitglieder abbilden. Diese betreffen die Ausbildung (Fachbereiche „Eduktion“ und „EbM im Studium“), Methodik (Fachbereich „Methodik“, „Health Technology-Assessment“), die praktische Anwendung (Fachbereich „EbM in Klinik und Praxis“, „Leitlinien“, „operative Fächer“, „Pfleger- und Gesundheitsförderung“ und „Zahnmedizin“) sowie übergreifend „Patienteninformation und Patientenbeteiligung“. Die Fachbereiche haben zwischen 36 und 267 Mitglieder. Sie haben zum Großteil eigenständige Aktivitäten entwickelt und eigene Produkte erarbeitet.

42.3 Leistungen

Tagungen

- ▲ Das DNEbM führt jährlich einen Kongress durch, auf dem aktuelle Themen und Ergebnisse aus dem Bereich Evidenzbasierte Medizin präsentiert und diskutiert werden können. Der Kongress dient gleichzeitig dem persönlichen Kontakt und dem Austausch der Netzwerkmitglie-

der, der ansonsten über elektronische Medien organisiert wird. Im Jahre 2006 wurde der Kongress zum ersten Mal gemeinsam in einer anderen wissenschaftlichen Fachgesellschaft, der Biometrischen Gesellschaft, durchgeführt, um so die Kontakte und Kooperationen auch außerhalb des Netzwerks auszubauen.

Stellungnahmen

- ▲ Das DNEbM hat in mehreren Stellungnahmen aktuelle Entwicklungen im Gesundheitssystem kommentiert, so zuletzt zum Thema Studienregister, zum Problem des „Probanden- und Patientenschutzes“, zu Fragen der „Regelungen zur Verbesserung der Qualität der Versorgung“ oder zur „Einführung eines PSA-Screenings“. Diese und weitere Stellungnahmen sind im Downloadbereich der Internetseite verfügbar.

Studie des Monats

- ▲ Seit 2005 wird in regelmäßigen Abständen auf der Internetseite eine „Studie des Monats“ vorgestellt und über eine Mailingliste von EbM-Interessierten werden methodische und medizinische Aspekte der Studie diskutiert. Ziel ist es hier unter anderem, eine kommentierte Beispielsammlung für EbM-Kurse zusammenzustellen.

Curriculum Evidenzbasierte Medizin

- ▲ Es wurde ein Curriculum gemeinsam mit dem Ärztlichen Zentrum für Qualität in der Medizin (ÄZQ) entwickelt, in dem Inhalte und Anforderungen für EbM-Kurse niedergelegt sind [3]. Die Anforderungen dieses Curriculums sind recht umfangreich, aber für einzelne Kurse sind Schwerpunktsetzungen erwünscht (s. Kap. 40).

Aus-, Fort- und Weiterbildung

- ▲ Es wurde die Möglichkeit für interessierte Anbieter geschaffen, ihre EbM-Kurse nach dem beschriebenen Curriculum zertifizieren zu lassen.
- ▲ Es wurden Train-the-trainer-Seminare zur Aus- und Fortbildung von Übungs- und Lehrkräften für EbM-Kurse durchgeführt.

Cochrane-Library

- ▲ Seit Anfang 2003 haben die Mitglieder des DNEbM kostenfreien Zugang auf die Cochrane-Library (s. Kap. 43), sodass ein Zugriff auf eine der wichtigsten Quellen für die praktische Umsetzung von EbM möglich ist. Hier können systematische Reviews, aber auch Originalarbeiten gesucht und für die praktische Arbeit verfügbar gemacht werden. Seit Anfang 2006 wird dieser Zugang durch eine Kooperation mit dem Deutschen Institut für Medizinische Dokumentation und Information (DIMDI) gewährleistet (s. Kap. 46).

Material

- ▲ Der so genannte „EbM-Splitter“ greift aktuelle Themen aus dem Bereich Evidenzbasierte Medizin auf und stellt sie einem deutschen Fachpublikum vor. Der EbM-Splitter erscheint als Rubrik in der ZaeFQ bzw. in der Deutschen Zahnärztlichen Zeitschrift. Bisher publizierte Beiträge können auf der Internetseite nachgelesen werden.
- ▲ In Kooperation mit dem DNEbM ist im Jahr 2003 das „Kompendium Evidenzbasierte Medizin als deutschsprachige Ausgabe von „Clinical Evidence Concise“,

erschienen [4]. Die zweite Auflage des Buches von Behrens/Langer „Evidence based Nursing“ wird ebenfalls in Anbindung an das DNEbM erscheinen.

- ▲ Seit 2004 steht den Mitgliedern des DNEbM auf der Internetseite ein Glossar mit wichtigen Fachbegriffen zur Verfügung (s. Kap. 49).

Das DNEbM wird sich in der Zukunft weiter für eine breite Anwendung Evidenzbasierter Medizin einsetzen. Eine wichtige Zielrichtung ist dabei, die EbM-Idee auch außerhalb der medizinischen Fachpresse zu verankern. Gleichzeitig stehen eine Förderung der notwendigen Kompetenzen innerhalb und außerhalb des Netzwerks, die Bearbeitung wissenschaftlicher Fragestellungen sowie eine Intensivierung der Kooperation mit medizinischen und methodischen Fachgesellschaften an.

Literatur

- [1] Jonitz G, Ollenschläger G, Kunz R, Evidence Based Medicine – Weg zu einer rationaleren Medizin. Dtsch Ärztebl (1998), 95, A 267–270
- [2] Kunz R et al. (Hrsg.) (2000) Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis. Deutscher Ärzte-Verlag, Köln
- [3] Antes G et al. (2005) Curriculum EbM – Ärztlicher Fortbildungskatalog Evidenzbasierte Medizin. DNEbM und ÄZQ, Berlin, <http://www.dnebm.de>
- [4] Ollenschläger G et al. (Hrsg.) (2003) Kompendium evidenzbasierte Medizin 03, Clinical Evidence Concise. Hans Huber, Bern Göttingen

43 Cochrane Collaboration: Deutsches Cochrane Zentrum/Deutsche Cochrane-Review-Gruppen

Antje Timmer, Gerd Antes

Die Cochrane Collaboration ist ein internationales Netzwerk, das sich die Erstellung, Verbreitung und regelmäßige Aktualisierung von systematischen Übersichtsarbeiten zur Aufgabe gemacht hat. Hintergrund ist die Erkenntnis, dass viele Ergebnisse klinischer Studien nicht für die Entscheidungsfindung am Krankenbett oder im Gesundheitswesen zur Verfügung stehen. Systematische Übersichten sollen hier durch ihre umfassende, systematisch und transparent begründete Würdigung hochwertiger Evidenz Abhilfe schaffen. Bisher konzentriert sich die Arbeit der Cochrane Collaboration auf die Identifikation und Synthese randomisiert-kontrollierter klinischer Studien zu Fragen der Intervention (Therapie und Prävention). Eine methodische Ausweitung zu nicht-randomisierten Studien und diagnostischen Fragestellungen zeichnet sich ab.

Grundsätzlich ist jeder Interessierte als Mitwirkender willkommen, ob Mediziner, sonstiger Mitarbeiter des Gesundheitssystems oder medizinischer Laie. Ansprechpartner für die Aktivitäten im deutschsprachigen Raum ist das Deutsche Cochrane Zentrum in Freiburg. Zudem sind zwei der weltweit zurzeit 51 Cochrane-Review-Gruppen in Deutschland angesiedelt. Diese Review-Gruppen sind themenbezogen vor allem für die Koordination der Review-Erstellung und -Begutachtung, daneben auch für die möglichst umfassende Erfassung klinischer Studien in ihrem Gebiet zuständig.

Die Cochrane Library ist das Produkt der Cochrane Collaboration, in der alle fertiggestellten Übersichtsarbeiten, aber auch alle Protokolle geplanter Projekte aufgenommen

sind. Zudem findet sich hier als weitere Ressource die wohl weltweit umfangreichste Datenbank zu kontrollierten klinischen Studien. Die Cochrane Library stellt damit eine ganz wesentliche Quelle für die systematische und umfassende Literatursuche in der EbM dar und ist z.B. für die Erstellung von Leitlinien und HTA-Berichten unverzichtbar.

Wir werden im Folgenden zunächst die Cochrane Collaboration und die Cochrane Library eingehender vorstellen, bevor wir im Einzelnen auf die Aktivitäten im deutschsprachigen Raum eingehen.

43.1 Die internationale Cochrane Collaboration

43.1.1 Entstehung und Zielsetzung

Namensgeber der Cochrane Collaboration ist der britische Epidemiologie Archibald Cochrane (1909–1988). Sein 1972 erschienenes Buch „Effectiveness and Efficiency. Random Reflections on the Health Services“ gilt als Schlüsselwerk und Initialzündung der Cochrane Collaboration [1]. Wichtig war ihm besonders die Bedeutung randomisierter Studien als Grundlage ärztlicher Entscheidungen. Die Kritik an der mangelnden Integration ihrer Ergebnisse aufgrund der Unüberschaubarkeit der Informationen nahm sich ein junger Assistenzarzt der Perinatalogie und Geburtshilfe zu Herzen: I. Chalmers begann 1974 mit dem Aufbau eines Studienregisters der Perinatalogie [2]. 1985 enthielt das Register bereits 3 500 Studien [3]. Bereits Ende der 70er-Jahre begann diese Arbeits-

gruppe mit der Erstellung systematischer Übersichten, die 1989 in dem wohl ersten im eigentlichen Sinne evidenzbasierten Lehrbuch veröffentlicht wurden [4]. Dieses Buch wird bis heute regelmäßig aufgelegt und ist auch in deutscher Übersetzung erhältlich [5].

Cochrane bezeichnete diese Sammlung von systematischen Übersichten im Vorwort zur ersten Ausgabe als „Meilenstein in der Geschichte randomisierter klinischer Studien und der Bewertung medizinischer Be-

handlungen“. Er forderte dabei die Vertreter aller klinischen Fächer auf, dem Beispiel der Perinatalogie zu folgen, erlebte dies selbst jedoch nicht mehr. 1992, fünf Jahre nach Cochranes Tod, wurde mit Unterstützung des National Health Service (NHS) das erste Cochrane-Zentrum in Oxford mit dem Ziel gegründet, die Erstellung systematischer Übersichtsarbeiten zu therapeutischen Fragestellungen auch in anderen Fächern voranzutreiben. Bereits im Jahr darauf wurde ausgehend von diesem ersten nationalen Zentrum die internationale Cochrane Collaboration gegründet.

Somit ist die Cochrane Collaboration ein gelungenes Beispiel einer Bottom-up-Organisation, bei der aus einer kleinen regionalen Initiative, dem Perinatalstudienregister in Oxford, ein weltweites Netzwerk entstand. Das Logo der Organisation zollt der Bedeutung dieses ersten Registers Tribut (s. Abb. 43.1).

So enthält es innerhalb der beiden C (für Cochrane Collaboration) die Ergebnisse einer Meta-Analyse, die zuerst im Geburtshilfebuch der Oxforder Gruppe publiziert wurde. Die vereinfacht dargestellte Graphik in

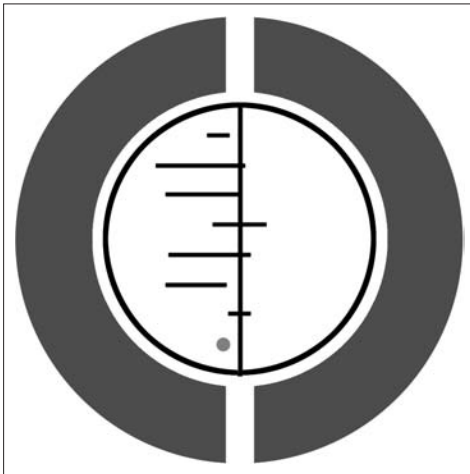


Abb. 43.1: Logo der Cochrane Collaboration

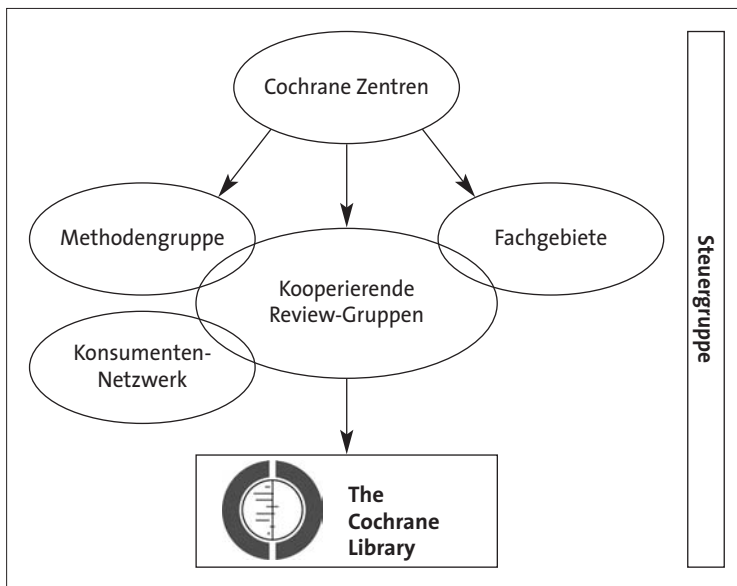


Abb. 43.2: Struktur der Cochrane Collaboration

Tab. 43.1: Grundsätze der Cochrane Collaboration

1. Zusammenarbeit durch leistungsfähige Kommunikationswege sowohl innerhalb der Cochrane Collaboration als auch nach außerhalb, durch transparente Entscheidungsstrukturen und durch Teamwork.
2. Enthusiasmus von Einzelpersonen als Basis der Netzwerkarbeit, Berücksichtigung und Unterstützung von Personen aus unterschiedlichen Fachgebieten oder Anwendungsbereichen.
3. Vermeidung von Doppelarbeit durch gutes Management und Koordination mit dem Ziel der größtmöglichen Effizienz.
4. Minimierung systematischer Fehler durch unterschiedliche Maßnahmen wie wissenschaftliche Strenge, Sicherstellung eines breit gefächerten Teilnehmerkreises und Vermeidung von Interessenkonflikten.
5. Aktualisierung durch die Verpflichtung von Autoren, Cochrane-Reviews durch Identifikation und Integration neuer Evidenz regelmäßig auf den neuesten Stand zu bringen.
6. Bestreben nach Relevanz durch die Förderung der Bewertung von gesundheitsbezogenen Interventionen nach Kriterien, die für die betroffenen Entscheidungsträger und Patienten wichtig sind.
7. Förderung der Zugänglichkeit durch die Verbreitung der Produkte der Cochrane Collaboration und Verhandlung angemessener Preise, Inhalte und Medien, um den Bedürfnissen weltweiter Nutzer zu entsprechen.
8. Sicherstellung der Qualität durch offenen Umgang mit Kritik, Berücksichtigung methodischer Fortschritte und Entwicklung von Systemen zur Qualitätsverbesserung.
9. Kontinuität durch Sicherung der fortgesetzten Verantwortlichkeit für Reviews, redaktionelle Betreuung und andere Schlüsselfunktionen.
10. Ermöglichen breiter Teilnahme an der Arbeit der Cochrane Collaboration durch Verminderung von Barrieren für Beiträge und durch Ermutigung zur Vielfalt.

Form eines Forest-Plots zeigt die Wirksamkeit von Steroiden bei vorzeitiger Wehentätigkeit (seither mehrfach aktualisiert [6]). Es ist im Übrigen auch ein trauriges Beispiel für die verspätete Umsetzung der Ergebnisse klinischer Studien. So hatte bereits das erste RCT von 1972 einen eindeutigen Effekt gezeigt, ab 1982 kann der Nachweis als gesichert gelten. Praktisch-klinische Beachtung erfuhr diese Erkenntnis jedoch erst seit der Publikation der Meta-Analyse ab 1989; inzwischen ist dies ein Standardverfahren.

Offizielle Mission der Cochrane Collaboration ist die Erstellung, regelmäßige Aktualisierung und Veröffentlichung von systematischen Übersichten zur Wirksamkeit von Interventionen im Gesundheitsbereich. Hierfür wurden zehn Grundsätze formuliert (vgl. Tab. 43.1) (www.cochrane.org).

43.1.2 Struktur der Cochrane Collaboration

Die Autoren eines Cochrane-Reviews („Reviewer“), können aus ganz unterschiedlichen Ländern und Fachrichtungen kommen. In der Regel tun sich multidisziplinäre Teams zusammen, sodass z.B. klinische Experten (Fachärzte), Methodiker der klinischen Epidemiologie, Recherche-Spezialisten und ggf. Statistiker gemeinsam an einem Review arbeiten. Die Kommunikationsstruktur der Cochrane Collaboration erlaubt und fördert eine internationale Zusammensetzung.

Diese Autorentams werden jeweils von einer themenspezifischen Cochrane-Reviewgruppe (CRG) betreut. Inzwischen gibt es 51 solcher Gruppen. Beispiele sind die aus der Oxforder Perinatalgruppe hervorgegangenen „Pregnancy and Childbirth“-Group, sowie die beiden unten dargestellten in Deutschland ansässigen Reviewgruppen

(Metabolische Erkrankungen und Hämatologie/Oncologie). Sie helfen bei der Zusammenstellung der Autorenteams, stellen Editoren und kümmern sich um die Berücksichtigung von Laien-Nutzer-Interessen. Die Cochrane-Review-Gruppen (CRGs) sind auch für die Erfassung klinischer Studien aus ihrem Themengebiet und die regelmäßige Aktualisierung eines spezifischen Studienregisters zuständig, welches dann in die Cochrane Library eingeht. Eine Handsuche beinhaltet z.B. die Durchsicht fachspezifischer Zeitschriften und Kongressbände auf klinische Studien [7].

Die CRGs werden ergänzt durch so genannte Fields, die sich krankheitsübergreifend für die adäquate Berücksichtigung besonderer Versorgungsaspekte, wie etwa der Komplementärmedizin, der Geriatrie oder der Rehabilitation, einsetzen. Eine besondere Bedeutung kommt dem Consumer Network zu. Es soll die Integration von Patienten und anderen Laien-Nutzern in die Arbeit der Cochrane Collaboration ermöglichen und fördern. Methodengruppen bearbeiten spezifische Fragestellungen zur Methodik systematischer Übersichtsarbeiten und sind damit besonders an der Weiterentwicklung der Cochrane Collaboration beteiligt. So befasst sich eine Methodengruppe mit der Bearbeitung diagnostischer Fragestellungen, die voraussichtlich in der bisher weitgehend auf interventionelle Studien beschränkten Cochrane Collaboration eine zunehmend größere Rolle spielen werden.

Alle Gruppen und Felder finden sich übersichtlich mit Kontaktdaten sowohl in der Cochrane Library (www.thecochrane.org) als auch auf den Seiten des Deutschen Cochrane Zentrums (www.cochrane.de) und den internationalen Seiten der Cochrane Collaboration (www.cochrane.org).

Cochrane-Zentren stellen das organisatorische Rückgrat der Cochrane Collaboration in den verschiedenen Ländern dar. So kümmern sie sich um den Auf- bzw. Ausbau der

Infrastruktur, einschließlich Schulung und Kommunikationsmöglichkeiten, die den Aktiven im Einzugsbereich die Arbeit in der Cochrane Collaboration ermöglichen sollen. Zurzeit gibt es weltweit zwölf, davon unterhalten einige zusätzliche Zweigstellen. Die verschiedenen Glieder der Cochrane Collaboration, wie Zentren, Felder, Gruppen und Consumer Network werden als Entitäten bezeichnet. Alle sind über gewählte Vertreter in der Steering-Group, dem politischen Kopf der Organisation, vertreten.

43.1.3 Systematische Übersichten der Cochrane Collaboration

Cochrane Reviews zeichnen sich durch einen besonders hohen Grad der Standardisierung aus. So durchlaufen bereits Protokolle einen editoriiellen Review-Prozess und werden in der Cochrane Library publiziert, wo sie öffentlicher Diskussion zugänglich sind. Besonders großer Wert wird auf eine umfassende Literatursuche gelegt, für die meist ein Rechercheexperte hinzugezogen wird. Für die Bearbeitung der Daten steht eine spezielle Software zur Verfügung. Dieser „Review-Manager“ kann kostenlos von den Webseiten der Cochrane Collaboration und ihrer Zentren heruntergeladen werden. Alle Übersichtsarbeiten werden mit einem Abstrakt, seit 2005 auch mit einer laienverständlichen Zusammenfassung, versehen, allerdings auf Englisch. Deutschsprachige Übersetzungen sind in Vorbereitung. Die Autoren verpflichten sich zu regelmäßigen Aktualisierungen, bei der inzwischen neu publizierte Studien nachträglich berücksichtigt werden.

Die Vorgehensweisen sind im Reviewer-Handbuch und Style-Guide, die ebenfalls frei verfügbar und öffentlich einsehbar sind, im Detail nachvollziehbar. Alle Schritte werden von der thematisch zuständigen Review-Gruppe betreut. Review-Gruppen unterhalten häufig auch umfangreiche Datenbanken auf

der Basis ausgefeilter thematischer Suchstrategien und Handsuchen, die für die Review-Erstellung herangezogen werden können.

Leider sind auch Cochrane Reviews nicht perfekt, sondern können in ihrer Qualität recht unterschiedlich sein. Doch konnte bereits 1992 gezeigt werden, dass sie im Durchschnitt Meta-Analysen anderer Quellen überlegen sind [8]. Die fortgesetzte Qualitätsverbesserung ist ein wesentliches Ziel, und weitere Fortschritte sind bereits erreicht worden [9]. Cochrane-Reviews stellen somit eine sehr wichtige Quelle für die unterschiedlichen Anwendungsbereiche der EbM dar.

43.2 Die Cochrane Library

Die Cochrane Library ist eine Datenbank, in die alle Produkte der verschiedenen Review-Gruppen Eingang finden. Sie ist online zugänglich und erscheint vierteljährlich. Seit 2005 wurde der Vertrieb vom Verlag Wiley übernommen, der für die Lizenzvergabe verantwortlich ist.

Die in der Cochrane Library enthaltenen Datenbanken können sämtlich kostenlos recherchiert werden. Auch die Zusammenfassungen und „Plain language summaries“ sind ohne Lizenz erhältlich.

Der Volltextzugriff ist kostenpflichtig. Viele Universitätskliniken und andere Einrichtungen haben eine Institutslizenz, über die ihre Mitglieder die Cochrane Library in vollem Umfang nutzen können. Auch verschiedene Gesundheitsportale und das Deutsche Netzwerk Evidenzbasierte Medizin bieten ihren Mitgliedern die Möglichkeit der Volltextnutzung der Cochrane Library. Die verschiedenen Zugangsmöglichkeiten haben wir auf www.cochrane.de übersichtlich zusammengefasst. Für Einzelpersonen ohne Zugang zu lizenzierten Instituten oder Abonnement ist in Deutschland für Einzelreviews derzeit die Volltextbestellung über DIMDI (www.dimdi.de) am kostengünstigsten.

Folgende Datenbanken sind in der Cochrane Library enthalten (Stand: Ausgabe 02/2006):

- ▲ Cochrane Database of Systematic Reviews (CDSR): Dies ist das eigentliche Produkt der Cochrane Library. Sie sind in der Cochrane Library unter „Cochrane Reviews“ zu finden. Dort bietet sich auch die Möglichkeit, über die Browse-Funktion nach Thema, Aktualität, zuständiger Review-Gruppe oder Alphabet sortierte Listen einzusehen. Enthalten sind zurzeit über 2500 abgeschlossene Arbeiten und knapp 1700 Protokolle.
- ▲ Database of Abstracts of Reviews of Effectiveness (DARE): Diese Datenbank ist in der Cochrane Library seit der Ausgabe 02/2006 unter „other reviews“ zu finden. Sie enthält Hinweise auf systematische Übersichten, die nicht im Rahmen der Cochrane Collaboration erstellt wurden, jedoch unter evidenzbasierten Gesichtspunkten evaluiert und abstrahiert wurden. Enthalten sind hier auch Arbeiten nicht-englischer Zeitschriften, die sonst möglicherweise nicht international zugänglich wären. Verantwortlich für die standardisierten Zusammenfassungen ist das Centre for Reviews and Dissemination (CRD) in York, Großbritannien. Da es sich nicht um Produkte der Cochrane Collaboration handelt, unterliegen die Zugriffsmöglichkeiten auf die Volltexte den Richtlinien der jeweiligen Herausgeber.
- ▲ Clinical Trials Register (CENTRAL; CCTR): Diese aktuell fast 500000 Referenzen führende Datenbank klinischer Studien ist in der Cochrane Library inzwischen unter dem Stichwort „clinical trials“ geführt. In CENTRAL gehen die Ergebnisse der Handsuchaktivitäten der verschiedenen Cochrane-Review-Gruppen und Zentren ein. Zudem werden regelmäßig Zitate kontrollierter klinischer Studien aus anderen Datenbanken, ins-

besondere MEDLINE und EMBASE, nach Überprüfung des Studiendesigns hochgeladen. CENTRAL gilt als das weltweit vollständigste Verzeichnis kontrollierter klinischer Studien und ist daher eine wesentliche Quelle bei der systematischen Evidenzsuche.

- ▲ Methodology Reviews und andere methodische Arbeiten: bis 2006 wurden Übersichtsarbeiten zu relevanten methodischen Aspekten der Cochrane Arbeit in einer eigenen Datenbank geführt. Diese sind seit April 2006 Bestandteil von CDSR. Sonstige Studien zur Methodik klinischer Studien und Meta-Analysen sind jedoch weiterhin in einer eigenständigen Datenbank vertreten, der Cochrane Review Methodology Database (CRMD), aktuell unter „methods studies“ mit 8255 Hinweisen in der Cochrane Library vertreten.
- ▲ Health Technology Assessments (HTA) und ökonomische Analysen (EE): Diese Datenbanken enthalten Hinweise auf HTA-Berichte und sonstige evidenzsynthetisierenden Analysen, wobei allerdings – anders als in CENTRAL und CDSR – kein Vollständigkeitsanspruch gegeben ist.

In der Cochrane Library finden sich außerdem verschiedene Hilfe- und Einführungsmaterialien, u.a. das Handbuch für Reviewer, ein Style-Guide für Cochrane Reviews, das Handbuch der Cochrane Collaboration und ein sehr umfangreiches Glossar. Anleitungen für die Nutzung der Cochrane Library werden in verschiedener Form angeboten, u.a. auch auf Deutsch.

Zur Suche in der Cochrane Library empfiehlt sich in der Regel die „advanced search“, bei der verschiedene Begriffe eingegrenzt und miteinander kombiniert werden können. Dabei ist es sinnvoll, zunächst die kurzen Hinweise zur Syntax anzuschauen. Sie werden bei der Nutzung der Advanced Search auto-

matisch eingeblendet. U.a. hinsichtlich Groß- und Kleinschreibung, Bindestrichnutzung und Trunkierung finden sich einige Besonderheiten. Hilfreich ist die starke Standardisierung der Review-Titel, die häufig eine rasche, auf Titelwörter beschränkte Suche ermöglicht. Alle Suchen werden jeweils in allen Datenbanken gleichzeitig durchgeführt.

Das Deutsche Cochrane Zentrum bietet regelmäßig Einführungen in die Cochrane Library an, sowohl im Rahmen von EbM-Grund- und Reviewerkursen als auch im Rahmen individueller, auch „In-House“-Schulungen. Diese Kurse werden auf der Webseite des Deutschen Cochrane Zentrums sowie über den Mail-Verteiler (s.u.) angekündigt.

43.3 Das Deutsche Cochrane Zentrum

43.3.1 Geschichte des Deutschen Cochrane Zentrums

Seit der Gründung der Cochrane Collaboration im Jahre 1993 verbreitete sich die Idee eines umfassenden Netzwerks rasch. Bereits im Gründungsjahr wurden nach dem UK-Zentrum in Oxford das Canadian, das Baltimore (heute US), und das Nordic Cochrane Zentrum, sowie die Neonatal-, Stroke-, Musculoskeletal- und Primary Health Care-Gruppen ins Leben gerufen. Gleichzeitig reagierte die US National Library of Medicine (NLM) mit der Einführung des Tags „Meta-Analysis“ und einer verbesserten Identifizierung randomisierter klinischer Studien in MEDLINE, und die EU mit der Bereitstellung von Geldern für die Identifikation von kontrollierten Studien aus europäischen allgemeinmedizinischen Zeitschriften. Dennoch vergingen weitere Jahre, bis die Idee auch in Deutschland Fuß fasste.

Ab 1995 begannen in Freiburg erste Cochrane-Initiativen, unterstützt vom Nie-

derländischen Cochrane Zentrum. So wurde im Rahmen des BIOMED 1 Programms der EU mit der Handsuche deutschsprachiger Zeitschriften begonnen. Dabei werden Zeitschriften auf Publikationen kontrollierter klinischer Studien überprüft [7; 10]. Zu den damals ersten untersuchten Zeitschriften gehörten z.B. die Deutsche Medizinische Wochenschrift und die Medizinische Welt. Inzwischen besteht mit CENTRAL eine Datenbank kontrollierter klinischer Studien, in die die Ergebnisse solcher Handsuchen eingehen, zusätzlich zum regelmäßigen Abgleich mit MEDLINE und anderen elektronischen Datenbanken. Ziel ist neben der Vermeidung von Publikations-Bias in systematischen Übersichten durch möglichst umfassende Berücksichtigung der gesamten vorhandenen Literatur zu einem Thema auch die bessere internationale Repräsentanz deutschsprachiger klinischer Forschung.

Die Protagonisten der Freiburger Cochrane-Initiative waren zeitgleich an der Förderung der Evidenzbasierten Medizin in Deutschland maßgeblich beteiligt. Aus diesen Erfahrungen heraus konnte nach entsprechender Antragstellung bei der Cochrane Collaboration ab 1998 das Deutsche Cochrane Zentrum in Freiburg seine Arbeit aufnehmen, initial gefördert durch das damalige Bundesministerium für Forschung und Technik. 2000 kamen zwei Cochrane-Review-Gruppen in Düsseldorf und Köln hinzu.

Anders als die Zentren in Großbritannien, Dänemark oder Australien ist das Deutsche Cochrane Zentrum bis heute kein konstant aus einheitlicher Quelle gefördertes Institut. Dennoch konnte es sich als eine feste und zuverlässige Größe sowohl in der internationalen Cochrane Collaboration als auch im Gesundheitssystem der Bundesrepublik und als Ansprechpartner des deutschsprachigen Raums etablieren. Zurzeit ist das Deutsche Cochrane Zentrum Teil des Instituts für Biometrie und Informatik am Universitätsklinikum Freiburg, welches die dauerhafte

Besetzung des Zentrum-Direktorats und des Sekretariats sicherstellt. Projektbezogene Finanzierungen ermöglichen die Beschäftigung weiterer Mitarbeiter. Über die nationalen Belange hinaus ist zudem das Webdevelopment-Team der Cochrane Collaboration am Deutschen Cochrane Zentrum untergebracht.

43.3.2 Aufgaben und Leistungen des Deutschen Cochrane Zentrums

Im Einzelnen ist das Deutsche Cochrane Zentrum für folgende Aufgabenbereiche zuständig:

43.3.2.1 Auf- und Ausbau der nationalen und internationalen Kommunikations- und Informationsstruktur

Das Deutsche Cochrane Zentrum ist der offizielle Vertreter der Cochrane-Aktivitäten aus Deutschland, Österreich, der deutschsprachigen Schweiz, Liechtenstein, Tschechien, Slowakei und Ungarn. Mitarbeiter des Deutschen Cochrane Zentrum sind regelmäßig und aktiv auf den internationalen Treffen und in den Arbeitsgruppen der Cochrane Collaboration vertreten.

Nachrichten aus der Cochrane Collaboration werden etwa halbjährlich über einen deutschsprachigen Rundbrief des Deutschen Cochrane Zentrums veröffentlicht. Zudem gibt es einen Mailverteiler, über den kurzfristig über wichtige Ereignisse und Neuigkeiten informiert werden kann. Diese Verteiler stehen allen Interessierten offen, ein Eintrag ist über mail@cochrane.de jederzeit möglich. Eine ständig aktualisierte Webseite soll den raschen und einfachen Zugriff auf alle wichtigen Informationen zur Arbeit der Cochrane Collaboration sowohl auf dem internationalen als auch dem nationalen Level ermöglichen (www.cochrane.de).

Aufgrund des im internationalen Vergleich verzögerten Starts der EbM-Aktivitätä-

ten in Deutschland hat sich das Deutsche Cochrane Zentrum an diversen Aktivitäten beteiligt, die oft deutlich über die Kernaufgaben eines Cochrane-Zentrums hinausgingen. Leitmotiv war und ist dabei immer, den Prinzipien der EbM und der klinischen Epidemiologie als ihrer methodischen Basis mehr Geltung zu verschaffen. Klar ist in Einklang mit der internationalen Sichtweise, dass der Kern der EbM „Orientierung an validen wissenschaftlichen Studien“ eine geeignete Wissensverdichtung erfordert, da der Rückgriff auf einzelne Studien für die Versorgungspraxis nicht durchführbar ist. Die Beteiligung am Aufbau der dafür notwendigen Strukturen hat das Deutsche Cochrane Zentrum immer als wesentliche Aufgabe angesehen. Struktur- und Curriculumentwicklung, Pilotprojekte und Beratungstätigkeiten sind stets darauf ausgerichtet, den Zugang zur international vorhandenen Evidenz in Deutschland und damit die Voraussetzungen für Evidenzberichte, HTA-Reports, klinische Leitlinien und Patienteninformation zu verbessern. Die Nachfrage nach Workshops und Schulungen hat in letzter Zeit stark zugenommen, da aufgrund der Verankerung der EbM im Sozialgesetzbuch und durch die Einrichtung des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) eine enorme Zunahme der Diskussion um EbM und ihrer Auswirkungen zu beobachten ist. Das Deutsche Cochrane Zentrum scheint vor allem aufgrund seines strikten Neutralitätskurses und aufgrund seiner internationalen Einbindung eine bevorzugte Anlaufstelle für diese Bedürfnisse zu sein.

43.3.2.2 Unterstützung der deutschen Reviewer-Tätigkeit

Primär sind die themenspezifischen Cochrane-Review-Gruppen für die Betreuung von Review-Autoren zuständig. Das Deutsche Cochrane Zentrum betrachtet es jedoch als eine vordringliche Aufgabe, die Reviewer-Tätigkeit im deutschsprachigen Raum voranzu-

treiben. Bei der Entwicklung der deutschen Webseiten wurde großer Wert darauf gelegt, die erforderlichen Ressourcen, Anleitungen und Kontaktadressen übersichtlich zur Verfügung zu stellen. Regelmäßig werden Schulungen zur Einführung in die Review-Erstellung sowie die Nutzung der Review-Software (Review-Manager) durchgeführt. Es stehen zudem Mitarbeiter für Rückfragen von potenziellen Review-Autoren jederzeit gerne zur Verfügung. So kann das Zentrum bei Vermittlung von Kontakten zu den zuständigen Gruppen helfen und methodische Beratung leisten. Mitarbeiter des Deutschen Cochrane Zentrums erstellen zudem auch selbst Reviews oder sind als Koautoren und zur methodischen Unterstützung an Reviews beteiligt, so z.B. in Zusammenarbeit mit der Hämatologie-Review-Gruppe in Köln oder verschiedenen klinischen Instituten und Abteilungen [11–15].

43.3.2.3 Aus- und Weiterbildung in Inhalten der EbM für Ärzte und Mitarbeiter im Gesundheitswesen

Mitarbeiter des Deutschen Cochrane Zentrums waren maßgeblich an der Entwicklung des Curriculums EbM (s. Kap. 40) beteiligt und sind deutschlandweit weiterhin in den Grund- und Aufbaukursen im Rahmen dieser Weiterbildung aktiv. Vom Deutschen Cochrane Zentrum werden regelmäßig (zweimal jährlich) Grundkurse mit dem Schwerpunkt RCT und systematische Übersichten durchgeführt, die außerordentlich hohen Zulauf haben. Zusätzlich werden nach Absprache z.B. Qualitätszirkel oder Krankenhäuser individuell geschult. Inhalte sind dabei z.B. die effiziente Literaturrecherche, die Arbeit mit der Cochrane Library oder die kritische Beurteilung von publizierten Studien. Das Deutsche Cochrane Zentrum bringt sich zudem in die Studentenausbildung in Freiburg ein und engagiert sich dafür, die traditionell vorwiegend methodisch orientierten biometrischen und epidemiologischen Inhalte im

Sinne einer mehr klinisch orientierten Anwendung zu vermitteln. Hier bietet die neue Approbationsordnung durch die Schaffung eines Querschnittsbereiches die geeignete Struktur, didaktische Erfahrungen aus der Evidenzbasierten Medizin (Ärzteschulungen) auch im Studentenunterricht umzusetzen.

43.3.2.4 Informationen für weitere

Zielgruppen: Laien, Journalisten und Entscheidungsträger im Gesundheitswesen

Ein wichtiger Schwerpunkt der Cochrane-Arbeit ist auch die Einbindung von Laien, die im Sprachgebrauch der Cochrane Collaboration aus deutscher Sicht unschön „Consumer“ genannt werden [16]. Dies wird umgesetzt durch die Einbindung im Rahmen des Consumer Networks z.B. als Kommentatoren von Reviews während des editoriiellen Prozesses. So können sie auf die Formulierung von Review-Protokollen Einfluss nehmen. Jede systematische Übersichtsarbeit enthält eine knappe, in laienverständlicher Sprache gehaltene Zusammenfassung, die so genannten Plain language summaries. Reisespenden sollen Betroffenen die Teilnahme an den Colloquien der Cochrane Collaboration erleichtern.

Eine Barriere für die Mitarbeit wie auch die Nutzung der Evidenz für Laien stellt sicherlich zurzeit noch die englische Sprache dar, die für eine aktive Teilnahme und Nutzung der Cochrane-Aktivitäten bisher unabdingbar ist. Inzwischen wurde mit der Übersetzung der laienverständlichen Zusammenfassungen begonnen. Für einen Großteil der Cochrane-Reviews dürften sie ab 2007 über die Webseite www.cochrane.de abrufbar sein. Eine ähnliche Initiative für spanischsprachige Länder hat im Vorjahr dort zu steilen Anstiegen der Zugriffszahlen geführt [17].

Das Deutsche Cochrane Zentrum war zudem an der Entwicklung des Curriculums Patientenberatung beteiligt. Unter der Projekt-

leitung des Ärztlichen Zentrums für Qualität in der Medizin (ÄZQ) wurde das Curriculum mit dem Ziel entwickelt, Patientenberatern (Laien) eine Schulung im Umgang mit Informationen, Recherche, Bewertung und Vermittlung anbieten zu können (s. Kap. 5, 18, 42). Dieses Curriculum ist unter www.patienten-information.de abrufbar. Erstmals wurde inzwischen ein spezieller Hands-on-Kurs für die laiengerechte Evidenzrecherche im Internet angeboten.

Zunehmend entwickelt sich auch die Zusammenarbeit mit Journalisten, sowohl im Rahmen von Journalistenschulungen zum Umgang mit Evidenz, als auch auf der Basis gemeinsamer Projekte zur Information der breiteren Bevölkerung. Ein Beispiel ist die 10-teilige Sternserie „Machen Sie sich schlau“ im Jahr 2006, mit dem Ziel der Förderung von Patientenkompetenz. Das Deutsche Cochrane Zentrum stellt zudem regelmäßig Experten zu aktuellen Themen der Gesundheitsinformation in den deutschen Medien.

43.3.2.5 Verbesserung der Repräsentanz deutscher Studien in der internationalen Wahrnehmung

Ständig fortgesetzt werden die Aktivitäten im Bereich der Handsuche, die 1995 im Rahmen des EU-Förderprogramms begonnen hatten. Inzwischen wurden etwa 17 000 klinische Studien aus 85 verschiedenen deutschsprachige Zeitschriften identifiziert und als Zitate in englischer Sprache international zur Verfügung gestellt [7]. Mehrere Publikationen sind bereits aus dieser Tätigkeit hervorgegangen, die den Umfang dieses Unternehmens für verschiedene Disziplinen verdeutlichen [18–20].

43.4 Die deutschen Cochrane-Review-Gruppen

43.4.1 Cochrane Metabolic and Endocrine Disorders Group, Düsseldorf

Die Cochrane Metabolic and Endocrine Disorders-Group (CMED) wurde als erste deutsche Cochrane-Gruppe im Februar 2000 registriert. Sie übernahm die Aufgaben der zuvor bestehenden Diabetes-Gruppe in Großbritannien, bei gleichzeitiger Erweiterung der Zuständigkeit auf Adipositas, sonstige Ernährungsstörungen und das gesamte Spektrum metabolischer und endokriner Erkrankungen. Kernanliegen von CMED ist die Sicherung hoher Qualitätsstandards der von der Cochrane Collaboration erarbeiteten systematischen Reviews. Derzeit sind über 200 verschiedene Projekte im editorielle Prozess von CMED.

Wer an einer aktiven Mitarbeit in diesem Themenbereich, z.B. als Review-Autor oder Consumer interessiert ist, sollte sich an diese Gruppe wenden, weitere Informationen finden sich unter <http://www.cc-endoc.info/>.

43.4.2 Cochrane Haematological Malignancies-Group (CHMG), Köln

Die Review-Gruppe der hämatologisch-onkologischen Erkrankungen wurde im Oktober 2000 als Teilprojekt des Kompetenznetzes Maligne Lymphome gegründet. Die der Klinik I für Innere Medizin der Universität zu Köln angegliederte Editorial Base wird zurzeit vom BMBF gefördert. Schwerpunktmäßig koordiniert diese Gruppe die Erstellung systematischer Übersichtsarbeiten im Gebiet der internistischen Hämato-Onkologie (vgl. Tab. 43.3).

Entsprechend der Konzeption von Cochrane-Review-Gruppen ist sie somit weltweit der Ansprechpartner für Personen, die in diesem Themengebiet eine Übersichtsarbeit erstellen möchten. Bisher wurden neun systematische Übersichten und 16 Protokolle fertiggestellt (Stand Juli 2006). Die Beteiligung der CHMG an nationalen und internationalen Projekten zur HTA-Berichterstattung und zur Entwicklung von Leitlinien (u.a. ASCO-ASH, AHRQ, DGHO, EORTC, IQWiG) unterstreicht die Versorgungsrelevanz der betreuten Themen.

Gefördert durch die DFG wurde in Zusammenarbeit mit der Deutschen Zentralbi-

Tab. 43.2: Erkrankungen und Verfahren im Zuständigkeitsbereich der CMED

Diabetes mellitus und assoziierte Erkrankungen	Störungen der Nebennieren
Diabetische Folgeerkrankungen (in enger Zusammenarbeit mit anderen CRGs)	Störungen der Schilddrüse
Störungen des Vitamin- und Lipidhaushaltes	Störungen der Hypophyse
Ernährungs-Defizienzen wie z.B. Proteinmangel	Adipositas
Sonstige metabolische Erkrankungen	Sonstige endokrine Erkrankungen

Tab. 43.3: Erkrankungen und Verfahren im Zuständigkeitsbereich der CHMG

Akute lymphatische Leukämie	Hodgkin Lymphome
Akute myeloische Leukämie	Non-Hodgkin Lymphome (NHL)
Chronische myeloische Leukämie	Multiples Myelom
Chronisch-lymphatische Leukämie	Myelodysplastisches Syndrom
Stammzelltransplantation	Aplastische Anämie

bibliothek für Medizin eine Datenbank klinischer Studien aufgebaut (www.haematology-studies.org). Ferner entstand in einem von der Deutschen Krebshilfe geförderten Projekt ein Patienten (Consumer-)Netzwerk (s. Kap. 41. – Schulungen von medizinischen Laien). Aktuelle Informationen zur CHMG finden sich unter www.chmg.de (Kontakt info@chmg.de).

43.4.3 Weitere Cochrane-Aktivitäten im deutschsprachigen Raum

Erfreulicherweise haben sich inzwischen über das Deutsche Cochrane Zentrum und die Review-Gruppen (Editorial Bases) hinaus einige Initiativen entwickelt, die aus dem deutschsprachigen Raum zur Arbeit der Cochrane Collaboration beitragen und von denen einige hier beispielhaft vorgestellt werden sollen.

So arbeitet die Arbeitsgruppe „Systematische Übersichtsarbeiten zur Behandlung von Persönlichkeitsstörungen“ an der Abteilung für Psychiatrie und Psychotherapie des Universitätsklinikums Freiburg eng mit der Cochrane Developmental, Psychosocial and Learning Problems-Group, Bristol/UK sowie dem Deutschen Cochrane Zentrum und seinem Heimatinstitut, der Abteilung für Biometrie und Statistik in Freiburg, zusammen. Aktuell werden Meta-Analysen zur psycho- und pharmakotherapeutischen Behandlung der Borderline-Persönlichkeitsstörung erstellt [21]. Diese Arbeitsgruppe unter Leitung von K. Lieb ist aufgrund langjähriger eigener Forschungstätigkeiten international mit Forschergruppen auf dem Gebiet der Behandlung von Persönlichkeitsstörungen vernetzt und verfügt so über einen guten Überblick über relevante aktuelle Forschungsaktivitäten.

Eine weitere Arbeitsgruppe mit psychiatrischem Schwerpunkt ist die Evidenz-basierte Psychiatrie an der TU München, die in-

zwischen eine ganze Reihe systematischer Übersichtsarbeiten im Bereich der Schizophreniebehandlung erstellt hat, z.B. [22; 23]. Der Leiter dieser Gruppe fungiert gleichzeitig als Editor der Cochrane Schizophrenia-Group. Im Kompetenznetzwerk Schizophrenie, das vom BMBF gefördert ist, ist die Gruppe als „Service Zentrum Cochrane Metaanalysen“ vertreten, sodass sie exemplarisch die Möglichkeiten der nationalen und internationalen Vernetzung auf der Grundlage einer lokalen Arbeitsgemeinschaft ausschöpft. Die Mitarbeiter engagieren sich auch in der Bearbeitung methodischer Fragestellungen der EbM, wie etwa der Qualität europäischer Guidelines im Bereich Psychiatrie, der klinischen Bedeutung der Ergebnisse psychiatrischer Rating-Skalen oder der Definition von „Response“ in Antipsychotika-Studien.

Durch eine Förderung der Deutschen Krebshilfe setzen sich onkologisch erfahrene Ärzte der Medizinischen Klinik 5/Institut für Hämatologie und Onkologie (Klinikum Nürnberg) in der Arbeitsgruppe Biologische Krebstherapie seit vielen Jahren intensiv mit unkonventionellen Verfahren in der Onkologie auseinander. Ziel dieser Arbeit ist, verlässliche Informationen zu unkonventionellen Verfahren in der Onkologie zu erarbeiten. Eine umfassende Literaturdatenbank wurde aufgebaut, die eine ideale Grundlage für die Aktivitäten in der Cochrane Collaboration bietet. Die AG Biologische Krebstherapie hat zwei Cochrane-Reviews, zur Misteltherapie und zum Einsatz von Selen, in verantwortlicher Autorenschaft fertiggestellt [24; 25], weitere Übersichten sind in Arbeit. Die Ergebnisse werden im Rahmen des EU-Projekts „CAM Cancer“, dessen Leitungsgremium die AG angehört, einer europaweiten Öffentlichkeit zur Verfügung gestellt (www.cam-cancer.org).

Bereits ab 1995 war K. Linde vom Zentrum für naturheilkundliche Forschung (ZnF) (II. Medizinische Klinik und Poliklinik

der TU München) in der Cochrane Collaboration engagiert. Das ZnF war maßgeblich an der Etablierung und Weiterentwicklung des Cochrane Complementary Medicine Field beteiligt. Insgesamt hat diese Gruppe bisher zehn Cochrane-Reviews erstellt, in der Mehrzahl in verantwortlicher Autorenschaft, z.B. [26; 27]. Auch diese Gruppe setzt sich mit methodischen Fragestellungen auseinander [28; 29].

In Wien ist die Universitätsklinik für Neurologie mit den Schwerpunkten Schmerzforschung, Botulinustoxin und Glaukomtherapie in der Cochrane Collaboration-Review-Arbeit tätig [30; 31]. Mehrere Arbeiten sind eingereicht oder in Vorbereitung.

Das Forschungszentrum für Balneologie und Kurortwissenschaft Bad Elster, eine nachgeordnete Forschungseinrichtung des Freistaates Sachsens, führt ebenfalls seit vielen Jahren systematische Reviews und klinische Studien zu nicht-pharmakologischen Behandlungen, vor allem im Bereich chronischer Erkrankungen, durch. In den letzten Jahren entwickelte Thomas Brockow dort federführend in Kooperation mit Kollegen der Universität München eine Review-Technik zum Outcome Assessment in klinischen Studien basierend auf der ICF (International Classification of Functioning, Disability and Health) [32; 33]. Der aktuelle Cochrane-Review des Instituts behandelt Sport bei Brustkrebs unter adjuvanter Therapie und erscheint in der Oktober-Ausgabe der Cochrane Library 2006.

Als letztes Beispiel sei die Arbeitsgruppe „Evidence-based Medicine and Clinical Economics“ der Klinik u. Poliklinik für Anästhesiologie der Universität Würzburg als aktive Initiative der Cochrane Collaboration in Deutschland genannt. Ziel ihrer Arbeit unter der Leitung von Dr. P. Kranke ist die Förderung einer evidenz-basierten perioperativen Medizin. Systematische Übersichtsarbeiten wurden bisher in den Gebieten Anästhesie, Schmerztherapie, Palliativmedizin und In-

tensivmedizin sowie in benachbarten und themenverwandten Gebieten verfasst, z.B. [34; 35].

Kontaktadressen und nähere Informationen zu diesen aktiven Gruppen finden sich auf den Webseiten des Deutschen Cochrane Zentrums (www.cochrane.de; „Aktivitäten im deutschsprachigen Raum“). Weitere aktive Consumer, Reviewer und Cochrane-Methodiker sind jederzeit herzlich eingeladen, dort ebenfalls über ihre Projekte zu berichten.

Literatur

- [1] Cochrane AL (1972) Effectiveness & Efficiency. Random Reflections on Health Services. The Nuffield Trust, London
- [2] King JF, A short history of evidence-based obstetric care. Best Pract Res Clin Obstet Gynaecol (2005), 19(1), 3–14
- [3] Chalmers I et al., The Oxford Database of Perinatal Trials: developing a register of published reports of controlled trials. Contr Clin Trials (1986), 7(4), 306–324
- [4] Chalmers I, Enkin M, Keirse M (1989) Effective Care in Pregnancy and Childbirth. Oxford University Press, Oxford
- [5] Enkin MW et al. (2005) Effektive Betreuung während Schwangerschaft und Geburt, 2, Hans Huber, Bern
- [6] Crowley PA, Antenatal corticosteroid therapy: a meta-analysis of the randomized trials, 1972 to 1994. Am J Obstet Gynecol (1995), 173(1), 322–335
- [7] Blümle A, 10 Jahre „Handsearching“ in Deutschland – ein Beitrag zum Cochrane Central Register of Controlled Trials. Z Arztl Fortbild Qualitätssich (2005), 99(6), 396
- [8] Jadad AR et al., Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. JAMA (1998), 280(3), 278–280
- [9] Shea B et al., A comparison of the quality of Cochrane reviews and systematic reviews published in paper-based journals. Eval Health Prof (2002), 25(1), 116–129
- [10] Hopewell S et al., A comparison of hand-searching versus MEDLINE searching to

- identify reports of randomized controlled trials. *Stat Med* (2002), 21(11), 1625–1634
- [11] Bohlius J et al., Recombinant human erythropoietins and cancer patients: updated meta-analysis of 57 studies including 9353 patients. *J Natl Cancer Inst* (2006), 98(10), 708–714
- [12] Bohlius J et al., Granulopoiesis-stimulating factors to prevent adverse effects in the treatment of malignant lymphoma. *Cochrane Database Syst Rev* (2004), 3:CD003189
- [13] Dietrich ES et al., Ceftriaxone versus other cephalosporins for perioperative antibiotic prophylaxis: a meta-analysis of 43 randomized controlled trials. *Chemotherapy* (2002), 48(1), 49–56
- [14] Knaebel HP et al., Systematic review and meta-analysis of technique for closure of the pancreatic remnant after distal pancreatectomy. *Br J Surg* (2005), 92(5), 539–546
- [15] Maier PC et al., Treatment of ocular hypertension and open angle glaucoma: meta-analysis of randomised controlled trials. *Br Med J* (2005), 331(7509), 134
- [16] Weingart O et al., Patientenbeteiligung in der Cochrane Collaboration – Barrieren, Erfahrungen und Konzepte aus deutscher Sicht. *Z Arzt Fortbild und Qualitätsich* (2005), 99(6), 367–371
- [17] Most accessed Cochrane reviews in The Cochrane Library and most accessed review abstracts on www.cochrane.org. Melbourne, Australia: 2005
- [18] Handsearching Trauma Surgery Journals – is it Worth the Effort? 8th European Conference of Medical and Health Libraries, Thinking globally – Acting locally. Cologne, Germany: ZBMed / EAHIL, 2002
- [19] Schulte J, Antes G, Türp J, Deutschsprachige Artikel zu kontrollierten Studien in der Mund-, Kiefer- und Gesichtschirurgie. *Schweiz Monatsschr Zahnmed* (2004), 114(3), 202–206
- [20] Galandi D, Schwarzer G, Antes G, The demise of the randomised controlled trial: Bibliometric study of the German-language health care literature, 1948 to 2004. *BMC Med Res Methodol* (2006), 6(1), 30
- [21] Lieb K et al., Borderline personality disorder. *Lancet* (2004), 364, 453–461
- [22] Leucht S, Hartung B. Perazine for schizophrenia. *Cochrane Database Syst Rev* (2006), 2, CD002832
- [23] Leucht S, Hartung B, Benperidol for schizophrenia. *Cochrane Database Syst Rev* (2005), 2, CD003083
- [24] Dennert G, Horneber M, Selenium for alleviating the side effects of chemotherapy, radiotherapy and surgery in cancer patients. *Cochrane Database Syst Rev* (2006), 3, CD005037
- [25] Horneber MA et al., Mistletoe Therapy in Oncology. *Cochrane Database Syst Rev* (2006), eingereicht
- [26] Linde K et al., Echinacea for preventing and treating the common cold. *Cochrane Database Syst Rev* (2006), 1, CD000530
- [27] Linde K et al., St John's wort for depression. *Cochrane Database Syst Rev* (2005), 2, CD000448
- [28] Linde K et al., Should systematic reviews include non-randomized and uncontrolled studies? The case of acupuncture for chronic headache. *J Clin Epidemiol* (2002), 55(1), 77–85
- [29] Linde K et al., Impact of study quality on outcome in placebo-controlled trials of homeopathy. *J Clin Epidemiol* (1999), 52(7), 631–636
- [30] Sycha T et al., Botulinum toxin in the treatment of rare head and neck pain syndromes: a systematic review of the literature. *J Neurol* (2004), 251 Suppl 1, I19–I30
- [31] Sycha T et al., Interventions for normal tension glaucoma. *Cochrane Database Syst Rev* (2003), 4, CD002222
- [32] Brockow T et al., Identifying the concepts contained in outcome measures of clinical trials on musculoskeletal disorders and chronic widespread pain using the International Classification of Functioning, Disability and Health as a reference. *J Rehabil Med* (2004), 44 Suppl, 30–36
- [33] Brockow T et al., Identifying the concepts contained in outcome measures of clinical trials on depressive disorders using the International Classification of Functioning, Disability and Health as a reference. *J Rehabil Med* (2004), 44 Suppl, 49–55
- [34] Kranke P et al., Pharmacological treatment of postoperative shivering: a quantitative systematic review of randomized controlled trials. *Anesth Analg* (2002), 94(2), 453–60, table
- [35] Kranke P et al., Hyperbaric oxygen therapy for chronic wounds. *Cochrane Database Syst Rev* (2004), 2, CD004123

44 Das ÄZQ – Kompetenzzentrum der ärztlichen Selbstverwaltung für Qualität und Evidenz in der Medizin

Günter Ollenschläger, Christian Thomeczek, Sylvia Säger, Monika Lelgemann

44.1 Einleitung

Die „Zentralstelle der deutschen Ärzteschaft zur Qualitätssicherung in der Medizin“ – Kurzform: „Ärztliches Zentrum für Qualität in der Medizin – ÄZQ“ (bis Anfang 2003: „Ärztliche Zentralstelle Qualitätssicherung“) ist eine gemeinsame Einrichtung von Bundesärztekammer und Kassenärztlicher Bundesvereinigung. Gegründet 1995, gehört das ÄZQ zu den Qualitätsinstitutionen, die in Trägerschaft oder Mit-Trägerschaft der ärztlichen Selbstverwaltungskörperschaften bundesweit agieren (s. Tab. 44.1).

Aufgabenschwerpunkte des ÄZQ sind:

- ▲ Sondierung und methodische Bearbeitung von Innovationen auf dem Gebiet der Qualitätsentwicklung in der Medizin,
- ▲ (Weiter-)Entwicklung der Methoden der evidenzbasierten Medizin,
- ▲ Entwicklung und Implementierung von Instrumenten der evidenzbasierten Medizin (z.B. Reviews, Leitlinien, Evidenzberichte, Patienteninformationen),

- ▲ Unterstützung der beiden Spitzenorganisationen der Ärzteschaft bei der Wahrnehmung ihrer Aufgaben auf dem Gebiet der Qualitätssicherung [1].

Im vorliegenden Kapitel werden Rahmenbedingungen, Arbeitsgebiete und -ergebnisse der Institution mit Bezug zur evidenzbasierten Medizin beschrieben.

44.2 Rahmenbedingungen

Eine wesentliche Eigenschaft des politischen Systems – und insbesondere des Gesundheitssystems – in Deutschland ist die Gewaltenteilung zwischen Bund, Ländern und legitimierten Organisationen der Zivilgesellschaft. Im Gesundheitssystem sind Kompetenzen an mitgliederbasierte, selbstregulierte, nicht-profitorientierte Organisationen von Leistungs- und Kostenträgern delegiert. Diese Selbstverwaltungsorganisationen, z.B. die Ärztekammern und Kassenärztlichen Vereinigungen,

Tab. 44.1: Qualitätsinstitutionen/institutionalisierte Qualitätsprogramme in (Mit-)Trägerschaft der ärztlichen Selbstverwaltung 1993–2006

Gründung	Name	Schwerpunkt	URL
1993	AQS	Prioritätensetzung, Informationsaustausch, Abstimmung zur Qualität im Gesundheitswesen aufgelöst 2004 (Aktivitäten teilweise übergeleitet in G-BA)	–
1995	ÄZQ	Methodik/Instrumente/Implementierung von EbM und QM (insbesondere Leitlinien, Patienteninformationen)	www.azq.de
2000	BQS	Externe Qualitätssicherung im stationären Bereich nach SGB V	www.bqs-online.de
2001	KTQ	Zertifizierung von Gesundheitseinrichtungen	www.ktq.de
2004	G-BA	Definition von Standards im Rahmen des SGB V	www.g-ba.de
2004	IQWiG	Evidenzbewertung (Zuarbeit zu G-BA)	www.iqwig.de

VI

nutzen Sachkenntnis und Motivation ihrer Mitglieder bei der Erfüllung ihrer Aufgaben. Die Einrichtungen der Selbstverwaltung basieren auf Pflichtmitgliedschaft und interner demokratischer Legitimierung [2] und werden von staatlichen Behörden beaufsichtigt.

Die Aufgaben der ärztlichen Körperschaften sind zwiespältig. Einerseits überträgt der Staat ihnen die Pflicht, auch im Sinne des Gemeinwohls für Ordnung im Beruf zu sorgen, notfalls mithilfe bestimmter Sanktionsmöglichkeiten (Berufsaufsicht). Andererseits haben die Körperschaften das Recht und die Pflicht, Politik und Staatsverwaltung in Fragen der Gesundheitspolitik fachlich zu beraten und – vor allem – die Belange des Berufsstandes insgesamt gegenüber Staat und Gesellschaft wahrzunehmen (Berufsrepräsentanz) [3]. Dabei obliegt der ärztlichen Selbstverwaltung insbesondere auch – gemeinsam mit Verbänden aus Wissenschaft und Praxis (Wissenschaftliche Fachgesellschaften und Berufsverbände) – die Definition, Kontrolle, Sicherung und Entwicklung der Qualität in der Medizin (s. Tab. 44.2).

Vor diesem Hintergrund und angesichts des zunehmenden Stellenwertes von Qualität und Transparenz des Gesundheitswesens in der öffentlichen Diskussion [4] empfahl der Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen 1989 die Einrichtung eines unabhängig agierenden, aber gemeinsam getragenen Instituts für die Förderung der Qualitätssicherung (QS) mit folgenden Aufgaben ([5] zitiert bei [6]):

- ▲ Forschungs- und Entwicklungsprojekte über QS,
- ▲ Entwicklung von Kriterien und Standards für die Beurteilung einer guten Versorgung,

- ▲ Hilfestellungen und Beratungen bei der lokalen Anwendung von QS,
- ▲ Schulung in QS,
- ▲ medizinische und ökonomische Evaluation der QS,
- ▲ Beurteilung der Qualität des Gesundheitswesens.

In der Folge kam es nicht zur Gründung eines einzelnen Instituts, sondern es wurden mehrere Organisationen mit unterschiedlichen Aufgabengebieten und Zielen etabliert (Übersicht siehe [7]).

Als das erste Qualitätsinstitut in gemeinsamer Trägerschaft der Spitzenorganisationen der ärztlichen Selbstverwaltung gründeten BÄK und KBV im März 1995 die Zentralstelle der deutschen Ärzteschaft zur Qualitätssicherung in der Medizin (ÄZQ) [8] (s. Tab. 44.1).

44.3 Organisationsform

Zwischen 1995 und 2002 entwickelte sich das ÄZQ von einem Koordinierungsbüro zwischen BÄK und KBV mit zunächst 4 Mitarbeitern zu einem Kompetenzzentrum für evidenzbasierte Medizin, Leitlinien und Patienteninformationen.

1997 wandelten die Gesellschafter BÄK und KBV das ÄZQ – unter Beibehaltung ihrer organisatorischen Rahmenbedingungen (s. Abb. 44.1) – in eine BGB-Gesellschaft um. Die Institution bearbeitet die ihr übertragenen Aufgaben mit folgenden Gremien und Einrichtungen: Verwaltungsrat, Planungsgruppe, Erweiterte Planungsgruppe, Geschäftsstelle und Expertenkreise.

Tab. 44.2: Aufgaben der ärztlichen Selbstverwaltung im Bereich Qualität der Gesundheitsversorgung

- Berufliche Qualifikation (Weiter- und Fortbildung)
- Definition und Verbreitung von Berufs- und Behandlungs-Standards (Leitlinien für Ärzte und Patienten)
- Förderung von Patientensicherheit und Behandlungsqualität (Qualitätsmanagement)
- Überwachung der Berufsnormen/ggfs. Sanktionierung
- Wissenschaftliche Beratung

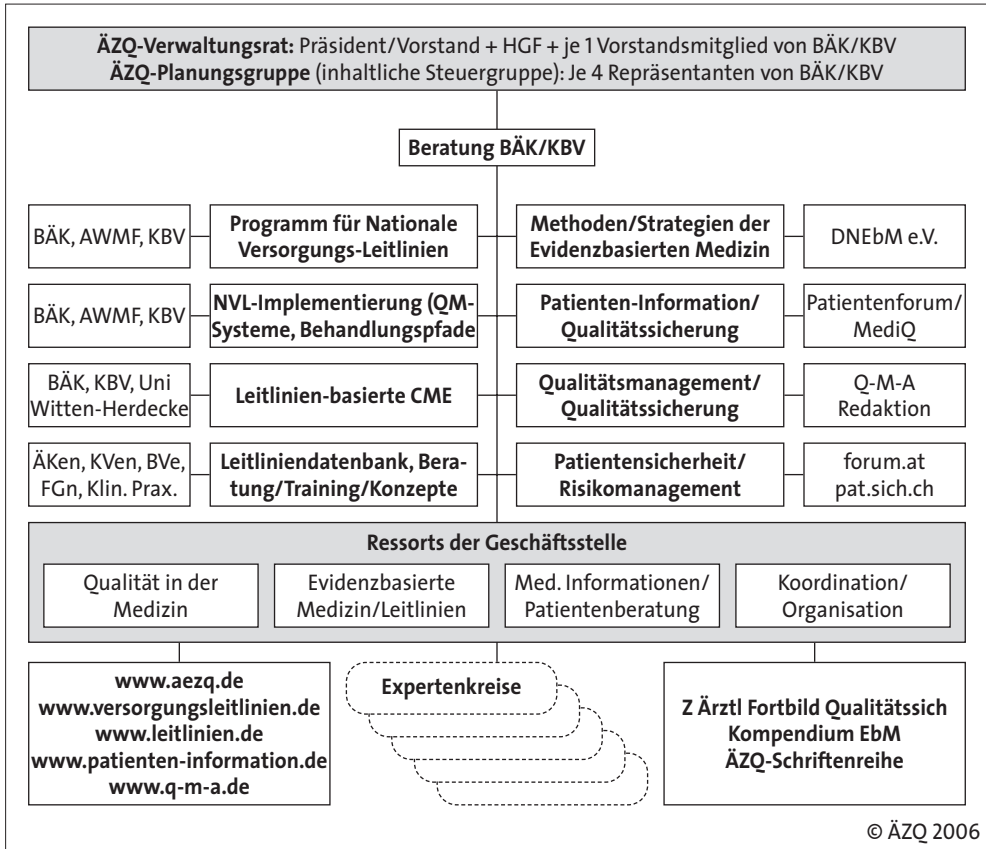


Abb. 44.1: Organigramm des ÄZQ mit Arbeitsgebieten und Kooperationspartnern (Stand: März 2006) [8]

Der **Verwaltungsrat** übt für die Gesellschafter die Rolle der Geschäftsführer des ÄZQ aus. Er beschließt im Rahmen der Vorgaben der Haushaltspläne beider Gesellschafter über den Haushaltsplan, die Erstellung der Jahresrechnung und die Finanzplanung. Der Verwaltungsrat trifft ferner die Entscheidungen in organisatorischen/personellen Grundsatzfragen. Der Verwaltungsrat beschließt einstimmig.

Der **Planungsgruppe** (= Vorstand) obliegt die fachliche Leitung der Institution. Sie trifft die grundsätzlichen Entscheidungen für die Arbeit und bestimmt die Arbeitsaufträge. Die Planungsgruppe entscheidet abschließend darüber, welche Entwürfe als Grundlage einer einheitlichen Beschlussfassung an die Organe der Gesellschafter weitergeleitet wer-

den. Mitglieder der Planungsgruppe sind je vier Repräsentanten der Träger. Die Planungsgruppe entscheidet einstimmig.

Eine so genannte **Erweiterte Planungsgruppe** (= Erweiterter Vorstand) wird für Fragen, welche im Rahmen der Beziehungen zu Drittorganisationen von Bedeutung sind eingerichtet. Eine erweiterte Planungsgruppe mit Repräsentanten der Träger und Kooperationspartner wurde sowohl im Rahmen des Deutschen Leitlinien-Clearingverfahrens als auch des Programms für Nationale Versorgungs-Leitlinien (s. Tab. 44.3) eingerichtet. Die Erweiterte Planungsgruppe entscheidet einstimmig.

Expertenkreise (Fachgruppen) werden von der Planungsgruppe für die Erarbeitung bestimmter Projekte berufen. Die Einrichtung von Expertenkreisen bedarf der Zustimmung

mung des Verwaltungsrats. Expertenkreise werden auf Zeit berufen. Die Planungsgruppe bestimmt die Vorsitzenden der Expertenkreise. Zwischen 1995 und 2005 unterstützten über 400 Experten aus allen Bereichen der medizinischen Wissenschaft und Praxis, aus Gesundheitswissenschaften, Gesundheitspolitik, klinischer Epidemiologie und Informationsmanagement das ÄZQ durch ehrenamtliche Arbeit. Darüber hinaus beteiligen sich Patienten und Konsumenten von Gesundheitsdienstleistungen an den Projekten.

Arbeitsgebiete und Arbeitsergebnisse 1995–2006

Die Existenz der ÄZQ zielt laut Gründungsvertrag von 1995 [8] auf:

- ▲ Koordination der Arbeit der ärztlichen Spitzenorganisationen auf dem Gebiet der Qualitätssicherung (s. Abb. 44.2),
- ▲ wirksame und einheitliche Entwicklung und Ausführung der Qualitätssicherung der ärztlichen Berufsausübung,
- ▲ Berücksichtigung der Interessen der Patienten,

- ▲ Kooperation mit Krankenkassen- und Krankenhausverbänden.

Vor diesem Hintergrund wurden die Aufgaben der Einrichtung im Vertrag von 1997 (Umwandlung in eine BGB-Gesellschaft) wie folgt formuliert:

1. Beratung/Unterstützung von BÄK/KBV bei ihren gemeinsamen Aufgaben der Qualitätssicherung der ärztlichen Berufsausübung.
2. Planung, Vorbereitung, Ausführung von Beschlüssen von BÄK/KBV.
3. Fachliche Vorbereitung für BÄK-Empfehlungen/KBV-Regelungen zur QS.
4. Unterstützung der Landesärztekammern und Kassenärztlichen Vereinigungen bei QS-Maßnahmen.
5. Beurteilung und Vorbereitung von wissenschaftlich begründeten und praktisch anwendbaren Leitlinien.

Auf dieser Grundlage wurden zwischen 1995 und 2006 etwa 100 Einzelprojekte realisiert (siehe [10]). Die wesentlichen Initiativen mit

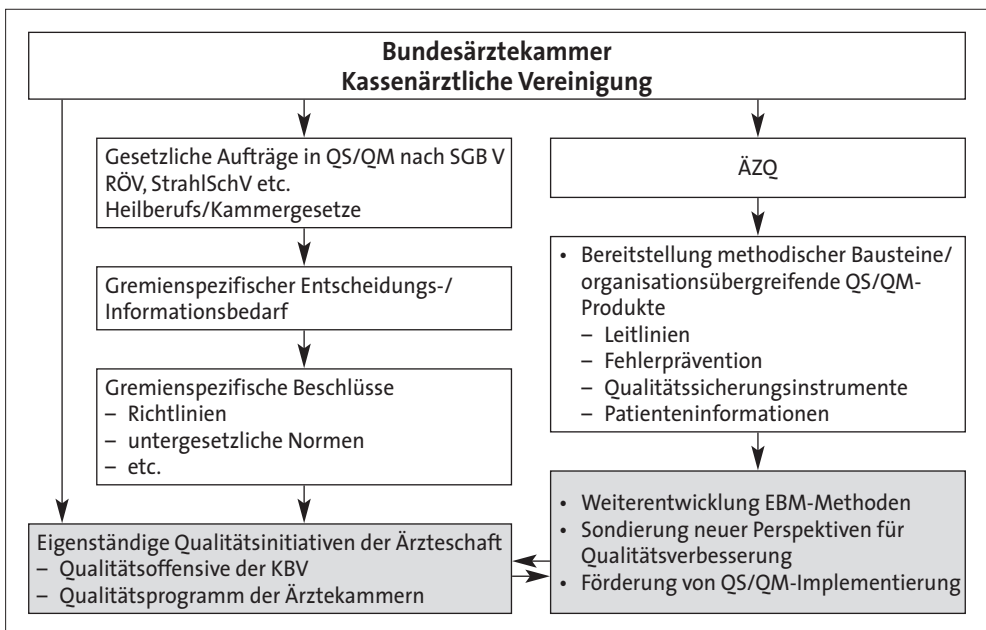


Abb. 44.2: Ziele und Aufgaben des ÄZQ als Kompetenzzentrum von BÄK und KBV [1]

Tab. 44.3: Initiativen von Bundesärztekammer und Kassenärztlicher Bundesvereinigung im ÄZQ zur Implementierung der evidenzbasierten Medizin unter Nutzung von Leitlinien 1995–2006

Jahr	Initiative	Referenz
1995	BÄK und KBV starten Leitlinien-Programm in der ÄZQ	[16]
	Bestandsaufnahme von BÄK und KBV zur Qualitätssicherung	[4]
	Schriftführerschaft der Zeitschrift für ärztliche Fortbildung und Qualität beim ÄZQ (später Organ von DNEbM, Cochrane-Zentrum, AWMF, ÄZQ, GIN, IQWiG)	[17]
1996	Beurteilungskriterien für Leitlinien von BÄK und KBV	[11]
1997	Entwicklung des Deutschen Leitlinien-Clearingverfahrens	[18–21]
	Leitlinien-Informationdienst www.leitlinien.de	[22]
1998	Initiative von BÄK und KBV zur Verbreitung von Leitlinien und evidenzbasierter Medizin gegenüber der Gesundheitsministerkonferenz	[12]
	Initiative zur Gründung des Deutschen Netzwerks Evidenzbasierte Medizin	[23]
1999	Institutionalisierung des Deutschen Leitlinien-Clearingverfahrens www.leitlinienclearing.de , Erster Leitlinien-Clearingbericht (Hypertonie)	[13; 24; 25]
	Koordination des Lehrbuchs Evidenzbasierte Medizin	[26]
2000	Leitlinienmanual von ÄZQ und AWMF www.leitlinienmanual.de	[27]
	Clearingverfahren für Patienteninformationen www.patienten-information.de	[14; 28; 29]
2001	Entwicklung der europäischen Standards für Leitlinien unter Beteiligung des ÄZQ	[30]
	Konzept zur Etablierung evidenzbasierter Gesundheitsziele	[31]
	Standards für evidenzbasierte Qualitätsindikatoren	[32]
2002	Programm für Nationale Versorgungs-Leitlinien der BÄK www.versorgungsleitlinien.de	[33]
	Schriftführerschaft des Kompendiums EbM beim ÄZQ	[34]
	Initiative zur Gründung des Leitlinien-Netzwerks G-I-N www.g-i-n.net	[35]
	Projekt zur Akzeptanz und Implementierung evidenzbasierter Leitlinien im hausärztlichen Bereich	[36]
2003	Initiative zur Nutzung von Leitlinien im Qualitätsmanagement www.q-m-a.de	[37; 38]
	Beitritt von AWMF und KBV zum Programm für Versorgungs-Leitlinien, Entwicklung einer gemeinsamen Methodik	[15; 39]
	Curriculum strukturierte medizinische Versorgung	[40]
2004	Leitlinienbasiertes Programm für interaktive, zertifizierte Fortbildung	[41]
	Curriculum für Patientenberater	[42]
2005	Deutsches Leitlinien-Bewertungsinstrument DELBI www.delbi.de	[43]
	Curriculum Evidenzbasierte Medizin	[44]
	Erste evidenzbasierte Nationale VersorgungsLeitlinie (Asthma)	[45]
	Kriterienkatalog zur Einschätzung der Qualität der medizinischen Versorgung durch Patienten (Checkliste Arztpraxis)	[46]
2006	Erste nationale Patientenleitlinie (Asthma)	[47]
	Handbuch zur Entwicklung regionaler evidenzbasierter Leitlinien	[48]
	Manual zur Erstellung evidenzbasierter Patientenleitlinien	[50]

Bezug zur evidenzbasierten Medizin sind in Tabelle 44.3 zusammengestellt.

Die Thematik „Evidenzbasierte Medizin“ wird im ÄZQ seit seiner Gründung vorrangig im Zusammenhang mit der Bewertung, Entwicklung und Implementierung von „wissenschaftlich begründeten und praktisch anwendbaren Leitlinien für die ärztliche Tätigkeit in der ambulanten und stationären Versorgung unter Berücksichtigung des Wirtschaftlichkeitsgebotes“ [8] bearbeitet.

Der Begriff „Evidenz“ selbst wird erstmals 1996 in den „Beurteilungskriterien für Leitlinien in der med. Versorgung“ verwendet [9; 11].

Durch

- ▲ die Identifizierung von Bundesärztekammer und Kassenärztlicher Bundesvereinigung mit dem Konzept der Evidenzbasierten Medizin (s. Tab. 44.4) [12],

- ▲ die Einrichtung der Clearingverfahren für Leitlinien (1999) [13] bzw. für Patiententinformationen (2000) [14] sowie
- ▲ die Institutionalisierung des Programms für Versorgungsleitlinien in gemeinsamer Trägerschaft von BÄK, AWMF und KBV – mit Unterstützung des Patientenforums (2003) [15]

konnten die Projekte des ÄZQ wesentlich zur Etablierung der Evidenzbasierten Gesundheitsversorgung in Deutschland beitragen, wie u.a. durch eine systematische Evaluation der Konsequenzen des Leitlinien-Clearingverfahrens belegt wurde (siehe Tab. 44.5. [13]).

Wichtige Voraussetzungen waren und sind dabei neben fachlicher Kompetenz und Engagement:

- ▲ der ehren- und hauptamtlich tätigen Experten des ÄZQ,

Tab 44.4: Empfehlung von Bundesärztekammer und Kassenärztlicher Bundesvereinigung zur Berücksichtigung der EbM in der Versorgung. Gemeinsame Stellungnahme gegenüber der GMK 1998 [12]

Nach Vorstellung der Ärzteschaft gilt es, die bereits in mehreren Ländern entwickelten und erprobten Strategien der Evidenzbasierten Medizin auf ihre Praktikabilität im hiesigen Versorgungssystem hin zu überprüfen.

Dabei betont die Ärzteschaft ausdrücklich, dass Evidenzbasierte Medizin im Sinne der ursprünglichen Definition von D. Sackett, Oxford, zu verstehen ist. Demnach bedeutet die Praxis der Evidenzbasierten Medizin (EbM), dass der praktizierende Arzt seine individuelle klinische Erfahrung mit der bestmöglichen, verfügbaren, externen klinischen Evidenz aus systematischer Forschung verbindet.

EbM zielt also auf optimale Nutzung der zur Verfügung stehenden wissenschaftlichen Erkenntnisse bei der diagnostisch-therapeutischen Entscheidungsfindung für den einzelnen Patienten. Evidenz ist dabei nicht auf randomisierte, kontrollierte Studien oder Metaanalysen beschränkt. Evidenz kann sich, je nach Problem, auch aus methodisch einwandfreiem Querschnitt/Follow-up-Studien, Leitlinien oder Ähnlichem ergeben. Des Weiteren ist zu prüfen, ob Ergebnisse der Evaluation von Modellvorhaben auch als Evidenz im Sinne des EbM-Paradigmas gelten können.

Die Vermittlung der aktuellen wissenschaftlichen Erkenntnisse an den praktizierenden Arzt bedarf eines mehrstufigen Vorgehens:

- Systematische Recherche und Bewertung der wissenschaftlichen Erkenntnisse – (Evidenz)-Methodik: Vorgehen der Cochrane-Collaboration;
- Aufbereitung der bewerteten Evidenz und Prioritätensetzung spezieller Handlungsempfehlungen – Methodik: Evidenz-basierte Konsensus-Leitlinien;
- Ständige Anpassung von Leitlinien und Empfehlungen an die fortschreitenden Entwicklungen der Medizin;
- Implementierung von Leitlinien mittels Aus-, Weiter-, Fortbildung, Qualitätszirkelarbeit, leitliniengestützter Dokumentationssysteme u.Ä.;
- Evaluation der ärztlichen Tätigkeit auf der Grundlage evidenzgestützter Leitlinien.

Tab. 44.5: Evaluation der Konsequenzen des Leitlinien-Clearingverfahrens 1999–2005 beim ÄZQ – Erreichte Ziele [13]

- Identifizierung/Darlegung der besten verfügbaren medizinischen Leitlinien
- Verbreitung der Methodologie und Nutzung evidenzbasierter Leitlinien
- Motivation der Fachgesellschaften zur Optimierung ihrer Leitlinien
- Implementierung des Leitliniengedankens im Bereich der Selbstverwaltung im Gesundheitswesen
- Etablierung eines nationalen Forums zum Austausch über evidenzbasierte Gesundheitsversorgung
- Internationale Platzierung der deutschen Leitlinienszene

- ▲ die enge Kooperation von BÄK und KBV im ÄZQ mit Partnern aus Selbsthilfe, Selbstverwaltung, Versorgungspraxis und Wissenschaft im In- und Ausland (z.B. Organisationen wie AkdÄ, AWMF, DKG, DNEbM, G-I-N, GKV, Patientenforum, PKV, RV und andere),
- ▲ die Etablierung des Deutschen Netzwerks Evidenzbasierte Medizin und die enge Einbindung des ÄZQ in seine Aktivitäten, Organisation, Publikationstätigkeiten. Internet: <http://www.aeqz.de>

im Workshop „Zukunft der Qualitätsentwicklung in der Medizin – Wie hilfreich sind neue Organisationsformen?“ veranstaltet von ÄZQ und AWMF in Berlin. http://www.aeqz.de/projekte/0index/projekt_abstracts/index/prj69_03/view (12.12.06)

- [7] Ollenschläger G (2006) Institutionalisierung der Qualitätsentwicklung in der Pflege und Betreuung. Gutachten für die Bundeskonferenz zur Qualitätssicherung im Gesundheits- und Pflegewesen (BUKO-QS). In Druck. BUKO-QS, Berlin
- [8] Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV) Vereinbarung der Bundesärztekammer und der Kassenärztlichen Bundesvereinigung über die Errichtung einer Zentralstelle der deutschen Ärzteschaft zur Qualitätssicherung in der Medizin. <http://www.aeqz.de/aeqz/0index/vertraege/view> (12.12.06)
- [9] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ) Tätigkeitsberichte 1996-2005. <http://www.aeqz.de/publikationen/taetigkeitsberichte/view> (12.12.06)
- [10] Ollenschläger G et al. (2003) ÄZQ Projekte 1995-2002. äzq Schriftenreihe; 10. videel, Niebüll
- [11] Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Beurteilungskriterien für Leitlinien in der medizinischen Versorgung – Beschlüsse der Vorstände der Bundesärztekammer und Kassenärztlicher Bundesvereinigung, Juni 1997. Dt Arztebl. (1997), 94, A-2154-5
- [12] Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV) Gemeinsame Stellungnahme zum Qualitätsmanagement im Gesundheitswesen gegenüber der Gesundheitsministerkonferenz. <http://www.aeqz.de/publikationen/stellungnahmen/0index/1998/view> (12.12.06)
- [13] Ielgemann M et al., Das Deutsche Leitlinien-Clearingverfahren 1999-2005-Hinter-

Literatur

- [1] Bundesärztekammer (BÄK) Tätigkeitsbericht 2004 der Bundesärztekammer. Dem 108. Deutschen Ärztetag 2005 in Berlin vorgelegt von Vorstand und Geschäftsführung. Kapitel 2. http://www.bundesaerztekammer.de/downloads/Taetigkeit2004_02.pdf
- [2] Busse R, Riesberg A, WHO Regional Office for Europe (2005) Gesundheitssysteme im Wandel. Deutschland. MWV, Berlin
- [3] Vogt G, Perspektiven der Selbstverwaltung im Gesundheitswesen. Rhein Arztebl. (2002), 14-15
- [4] Ollenschläger G (1996) Qualität und Qualitätssicherung in der Medizin. Grundlagen einer bedarfsgerechten Gesundheitsversorgung. Gemeinsame Bestandsaufnahme von BÄK und KBV. Zuckschwerdt, München
- [5] Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen (1989) Jahrgutachten 1989: Qualität, Wirtschaftlichkeit und Perspektiven der Gesundheitsversorgung. Nomos, Baden-Baden
- [6] Selbmann HK, Qualität im deutschen Gesundheitswesen – Wo stehen wir? Vortrag

- grund, Zielsetzung, Ergebnisse. Abschlussbericht. <http://www.leitlinien.de/clearingverfahren/index/clearing/view> (28.06.06)
- [14] Sanger S (2004) Einbeziehung von Patienten/Verbrauchern in den Prozess des Qualitatsmanagements im Gesundheitswesen am Beispiel der Qualitatsforderung medizinischer Laieninformation im Internet. azq Schriftenreihe; 17. videel, Niebull
- [15] Kopp I, Lelgemann M, Ollenschlager G, Nationale VersorgungsLeitlinie Asthma: Methodik und Entwicklungsprozess. Z Arztl Fortbild Qualitatssich. (2006), 100, im Druck
- [16] Ollenschlager G, Thomeczek C, Arztliche Leitlinien-Definitionen, Ziele, Implementierung. Z Arztl Fortbild Qualitatssich. (1996), 90, 355–361
- [17] Ollenschlager G, Geidel H, Klein-Lange M, Editorial zum neuen Herausgabeteam der Zeitschrift fur arztliche Fortbildung. Z Arztl Fortbild (Jena). (1995), 89, 3–4
- [18] Lauterbach KW et al., Konzept eines Clearingverfahrens fur Leitlinien in Deutschland. Z Arztl Fortbild Qualitatssich. (1997), 91, 283–288
- [19] Ollenschlager G et al., Die Checkliste zur methodischen Qualitat von Leitlinien. Ein Beitrag zur Qualitatsforderung arztlicher Leitlinien. Z Arztl Fortbild Qualitatssich. (1998), 92, 191–194
- [20] Antes G et al., Die methodische Qualitat von Leitlinien – dargestellt am Beispiel „Asthma bronchiale“. Z Arztl Fortbild Qualitatssich. (1998), 92, 295–297
- [21] Helou A, Ollenschlager G, Ziele, Moglichkeiten und Grenzen der Qualitatsbewertung von Leitlinien. Ein Hintergrundbericht zum Nutzermanual der Checkliste „Methodische Qualitat von Leitlinien“. Z Arztl Fortbild Qualitatssich. (1998), 92, 361–365
- [22] Lampert U et al., Der Online-Informationssdienst „LEITLINIEN-IN-FO“. Ein Beitrag zum Qualitatsmanagement im Gesundheitswesen. Z Arztl Fortbild Qualitatssich. (1999), 93, 39–44
- [23] Jonitz G, Evidence Based Medicine. Weg zu einer rationaleren Medizin. Dt Arztebl. (1998), 95, A–267–70
- [24] Ollenschlager G et al., Das Leitlinien-Clearingprogramm der Selbstverwaltungskorperschaften im Gesundheitswesen. Ein Projekt zur Qualitatsforderung in der Medizin. Gesundheitswesen. (1999), 61, 105–111
- [25] Donner-Banzhoff N et al. (2000) Leitlinien-Clearingbericht „Hypertonie“. azq Schriftenreihe; 5. Zuckschwerdt, Munchen
- [26] Kunz R, Ollenschlager G, Raspe H (2000) Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis. Dt. Arzte-Verlag, Koln
- [27] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Arztliche Zentralstelle Qualitatssicherung (AZQ), Das Leitlinien-Manual von AWMF und AZQ. Z Arztl Fortbild Qualitatssich. (2001), 95 Suppl 1, 1–84
- [28] Charnock D et al. (2000) Das DISCERN-Handbuch. Qualitatskriterien fur Patienteninformationen uber Behandlungsalternativen. Nutzerleitfaden und Schulungsmittel. azq Schriftenreihe, Band 3. Zuckschwerdt, Munchen, <http://www.discern.de>
- [29] Sanger S et al., Gut informiert uber Gesundheitsfragen, aber wie? Das Deutsche Clearingverfahren fur Patienteninformationen – Zielsetzung, Hintergrund und Arbeitsweise. Gesundheitswesen. (2002), 64, 391–397
- [30] Europarat, Entwicklung einer Methodik fur die Ausarbeitung von Leitlinien fur optimale medizinische Praxis. Empfehlung Rec(2001)13 des Europarates und Erluterndes Memorandum. Deutschsprachige Ausgabe. Z Arztl Fortbild Qualitatssich. (2002), 96 Suppl III, 1–60
- [31] Ollenschlager G (2002) Tragfahige Gesundheitsziele-Orientierung an wissenschaftlichen Erkenntnissen. GVG Schriftenreihe; 37. In: Gesellschaft fur Versicherungswissenschaft und -gestaltung (GVG), Gesundheitsziele.de-Forum Gesundheitsziele Deutschland: Gesundheitsziele fur Deutschland; Entwicklung, Ausrichtung, Konzepte, 51–69. AKA, Berlin
- [32] Altenhofen L et al., Beurteilung klinischer Messgroen des Qualitatsmanagements-Qualitatskriterien und -indikatoren in der Gesundheitsversorgung. Konsenspapier von BAK, KBV und AWMF. Z Arztl Fortbild Qualitatssich. (2002), 96, 2–15
- [33] Ollenschlager G, Diabetes mellitus: Erste Nationale Versorgungsleitlinie erschienen. Dtsch Arztebl. (2002), 99, A–1485–6
- [34] Ollenschlager G et al. (2003) Kompendium evidenzbasierte Medizin 2003. Clinical Evidence Concise. Huber, Bern

- [35] Ollenschläger G et al., Improving the quality of health care: using international collaboration to inform guideline programmes by founding the Guidelines International Network (G-I-N). *Qual Saf Health Care.* (2004), 13, 455–460
- [36] Kirchner H (2003) Das Deutsche Leitlinien-Clearingverfahren. Hintergrund, Zielsetzung, Ergebnisse dargestellt an Leitlinien zur Behandlung des Tumorschmerzes. Inaugural-Dissertation zur Erlangung der Doktorwürde. äzq Schriftenreihe; 13. Universitätsklinikum, Köln
- [37] Gramsch E et al. (2003) Kompendium Q-M-A. Qualitätsmanagement in der ambulanten Versorgung. Dt. Ärzte-Verlag, Köln
- [38] Thomeczek C et al., Die Checkliste Q-M-A. Ein Instrument zur Bewertung von Qualitätsmanagementsystemen in der ambulanten Versorgung. *Gesundheitswesen.* (2003), 65, 585–592
- [39] Ollenschläger G et al., Medizinische Leitlinien in Deutschland, 1994 bis 2004. Von der Leitlinienmethodik zur Leitlinienimplementierung. *Z Arztl Fortbild Qualitatssich.* (2005), 99, 7–13
- [40] Boesen N et al. (2003) Curriculum strukturierte medizinische Versorgung. Texte und Materialien der BÄK zur Fortbildung und Weiterbildung; 24., Köln
- [41] Vollmar HC et al., Realitätsnah lernen. Die Fortbildungsplattform Leitlinien-Wissen.de soll dazu beitragen, evidenz- und leitlinienbasiertes Wissen in die Praxis zu tragen und das lebenslange Lernen zu fördern. *Dt Arztebl.* (2004), 101, A–3480–2
- [42] Dierks ML et al., Curriculum Patientenberatung. Informationen recherchieren, bewerten, vermitteln. 2nd ed. http://www.patienten-information.de/content/projek-te/kontinuierlich/download/curriculum_ebm_patinfo_akt.pdf (12.12.06)
- [43] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI). Fassung 2005/2006. *Z Arztl Fortbild Qualitatssich.* (2005), 99, 468–519
- [44] Antes G et al. (2005) Curriculum EbM-Ärztlicher Fortbildungskatalog Evidenzbasierte Medizin., Berlin
- [45] Hoppe J et al., Nationale Versorgungsleitlinie Asthma. *Dt Arztebl.* (2005), 102, A–2734–9
- [46] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ) Woran erkennt man eine gute Arztpraxis? Checkliste für Patientinnen und Patienten. äzq Schriftenreihe; 23. http://www.patienten-information.de/content/informationsqualitaet/checkliste_arztbesuch/view (08.12.05)
- [47] Bundesärztekammer (BÄK), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Kassenärztliche Bundesvereinigung (KBV). Patientenleitlinie Asthma. <http://www.versorgungsleitlinien.de/patienten/asthmainfo> (22.05.06)
- [48] Schubert I et al. (2006) Handbuch zur Entwicklung regionaler Leitlinien. äzq Schriftenreihe; 26. BoD, Norderstedt
- [50] Sängler S et al., Manual Patienteninformation. Empfehlungen zur Erstellung evidenzbasierter Patienteninformationen. äzq Schriftenreihe; 25. http://www.patienten-information.de/content/download/manual_patienteninformation_04_06.pdf (23.05.06)

45 Das Netzwerk der Koordinierungszentren für Klinische Studien

Christoph M. Seiler, Insa Bruns, Stephanie Wolff, Walter Lehmacher, Markus Löffler, Herbert Maier-Lenz

45.1 Hintergrund

Die Initiative der Bundesregierung zur Gründung der Koordinierungszentren für Klinische Studien (KKS) geht bis in die Mitte der 1990er-Jahre zurück. Zu diesem Zeitpunkt zeichnete sich die Forschungslandschaft in Deutschland zwar durch eine gute Grundlagen- sowie krankheitsorientierte Forschung aus, es mangelte jedoch an patientenorientierten Forschungsansätzen. Der hohe Aufwand, die lange Dauer klinischer Studienprojekte, das vergleichsweise mäßige Ansehen von multizentrischen Studien als Forschungsleistung und die Überschätzung der praktischen Anwendbarkeit der Ergebnisse der biomedizinischen Grundlagenforschung werden als Gründe für die geringen Forschungsaktivitäten in Deutschland im Vergleich zu anderen Ländern aufgeführt. Das Ergebnis wird durch die Analyse von M. Rothmund, Marburg, aus dem Jahre 1997 bestätigt [1]: So kamen von 637 in den Zeitschriften *Lancet* und *New England Journal of Medicine* in den Jahren 1993 bis 1997 publizierten randomisierten klinischen Studien 243 aus den USA, 92 aus Großbritannien, 43 aus den Niederlanden, 33 aus Frankreich, 29 aus Italien und nur 19 aus Deutschland. Um zur Verbesserung dieser Lage beizutragen und eine von den Gesichtspunkten der Evidenzbasierten Medizin (EbM) geprägte Studienkultur zu fördern, gab das Bundesministerium für Bildung und Forschung (BMBF) den entscheidenden Impuls zur Gründung von studienbezogenen Wissenschaftseinrichtungen in Deutschland: Im Rahmen einer kompetitiven Ausschreibung des BMBF wurden im Jahr 1999/2000 und in

einer zweiten Ausschreibung im Jahr 2002 insgesamt 12 Koordinierungszentren für Klinische Studien (KKS) an Medizinischen Fakultäten gebildet [2].

45.1.1 Ziele und Auftrag des KKS-Netzwerks (KKS-N)

Die Mitglieder des KKS-Netzwerks (KKS-N) haben das Ziel, die patientenorientierte klinische Studienforschung in Deutschland nachhaltig zu verbessern, klinische Studien aus dem akademischen Umfeld (nicht-kommerzielle Studien) und aus der Wirtschaft zusammen mit den Institutionen der Krankenversorgung durchzuführen und die erforderlichen Methoden zu entwickeln. Die Mitglieder haben sich dabei dem Ziel verpflichtet, solche Studien nach dem Stand der Wissenschaft und den erforderlichen Qualitätsstandards durchzuführen.

Dabei bietet das KKS-N regional und überregional ein breites Spektrum an Wissenschaftsdienstleistungen. Studien mit Kindern und Jugendlichen werden angegliedert in spezifischen pädiatrischen Modulen (Freiburg, Heidelberg, Köln, Leipzig, Mainz und Münster) betreut (www.paed-net.org). Durch die Kooperation mit dem Studienzentrum der Deutschen Gesellschaft für Chirurgie (SDGC) in Heidelberg und den chirurgischen Regionalzentren (Berlin, Kiel-Lübeck, Köln, Mainz, München) können chirurgische Studien realisiert werden. Perspektivisch ist eine Angliederung des chirurgischen Netzwerks entsprechend der Rolle des pädiatrischen Netzwerks in das KKS-N geplant.

Die Mitglieder arbeiten gemeinsam daran, das wissenschaftliche Profil zu fördern und damit die klinische Studienforschung als wissenschaftliche Disziplin in Deutschland weiterzuentwickeln. An den Mitglieds-einrichtungen wird eigenständige medizinische und methodische Forschung betrieben. Das KKSN fördert den fachlichen Austausch hierzu durch geeignete wissenschaftliche Veranstaltungen und die Kooperation in Verbundforschungsprojekten. Auch engagiert sich das Netzwerk auf strategisch wichtigen Zukunftsfeldern, wie bei methodischen Forschungsprojekten sowie Studien im Bereich der klinischen Genomik und im versorgungsnahen Bereich (sog. „pragmatic trials“). Dabei unterstützt das Netzwerk insbesondere nicht-kommerzielle Studienprojekte im Bereich der Versorgungsforschung, sofern es sich um interventionelle klinische Studien mit einem vergleichenden Design handelt.

Die Koordinierungszentren sehen enge Beziehungen zwischen ihren Aktivitäten und dem Konzept der EbM. Innovative Studien zu Fragestellungen, für die aufgrund von vorhandenen oder noch zu erstellenden systematischen Reviews besonderer Forschungsbedarf besteht, sollen initiiert und unterstützt werden. Andererseits soll die aktive Beteiligung an Review-Gruppen in Zusammenarbeit mit den entsprechenden Fachgesellschaften und dem Deutschen Cochrane Zentrum in Freiburg durch das KKSN gefördert werden. Mit den KKS steht eine Infrastruktur zur Verfügung, die für Studien von öffentlichem Interesse, wie beispielsweise Head-to-head-Studien, genutzt werden kann – eine Finanzierung durch die öffentliche Hand vorausgesetzt.

Über die netzwerkinterne Zusammenarbeit hinaus kooperiert das KKSN mit medizinischen Fachgesellschaften, dem Deutschen Cochrane Zentrum, dem Deutschen Netzwerk Evidenzbasierte Medizin, den Kompetenznetzen in der Medizin, der Telematikplattform für Medizinische Forschungsnetze

(TMF e. V.) und anderen, auch europäischen Netzwerken (beispielsweise European Clinical Research Infrastructure Network). Darüber hinaus gibt es eine Vielzahl von Kooperationen mit lokalen sowie überregionalen, auch europäischen Kooperationspartnern und Studiengruppen. Letztere sind mit mehreren hundert Prüfzentren in bis zu sechs Ländern verbunden.

Prozesse der IT-basierten Datenerfassung, Studienorganisation, Projektmanagement, Qualitätssicherung, Datenweitergabe und Auswertung werden im Netzwerk harmonisiert und weiterentwickelt. Dies schließt perspektivisch die Entwicklung eines übergeordneten Konzeptes zur Integration vorhandener Daten und zur Modellierung von Prozessen ein.

45.1.2 Aufbau und Struktur

Im Rahmen der BMBF-Fördermaßnahme wurden von 1999 bis 2002 zwölf Koordinierungszentren für Klinische Studien gegründet. Diese sind an medizinische Fakultäten angebunden. Die Finanzierungsstruktur umfasst eine Grundausrüstung durch die lokale Fakultät, eine degressive Förderung durch das BMBF (75% in den Jahren 1 und 2, 50% in den Jahren 3 und 4, 25% in den Jahren 5 und 6) und zunehmende Einkünfte aus öffentlichen und Industriemitteln. Die Organisationsform sieht in der Regel einen Vorstand mit einem Wissenschaftlichen Leiter und einen Geschäftsführer vor. Die KKS handeln an ihrem Standort eigenständig.

Jedes KKS ist in ein inneruniversitäres Netzwerk mit breiter klinischer und methodischer Expertise eingebunden. Durch die Nähe zu den heimischen Instituten kann damit ein großer Wissensbereich abgedeckt werden (z.B. Biometrie, Klinische Pharmakologie).

307 Mitarbeiter sorgen deutschlandweit für professionellen Support im Umfeld klini-

scher Studien. Zu den Netzwerk-Experten zählen 38 Studienkoordinatoren (Ärzte), 25 Biometriker, 42 so genannte Clinical Research Associates, 45 Studienassistenten, 56 Datenmanager, 30 IT-Fachleute und 12 Rotationsärzte (Stand: 8/2006). Zusätzliche personelle Ressourcen werden durch mit dem Netzwerk kooperierende Kliniken und Institute bereitgestellt.

Die einzelnen Standorte haben von Beginn an im Rahmen einer Arbeitsgemeinschaft eng zusammengearbeitet. 2005 wurde diese Arbeitsgemeinschaft formal in ein Konsortium umgewandelt. Oberstes Organ des Konsortiums ist die Mitgliederversammlung. Die Geschäfte des Konsortiums werden durch einen dreiköpfigen Vorstand geführt. Die enge Kooperation zwischen den einzelnen KKS wird durch die gemeinsame Geschäftsstelle für übergreifende Aufgaben unterstützt. In den Fachgruppen „Aus-, Fort- und Weiterbildung“, „Qualitätsmanagement“, „Datenmanagement/Biometrie“ sowie themenbezogenen Task Forces werden gemeinsame Grundsätze und Vorgehensweisen erarbeitet. Die KKS werden regelmäßig Systemaudits im Auftrag des BMBF unterzogen.

Zum KKS-Netzwerk gehören heute neben den zwölf Koordinierungszentren für Klinische Studien (vergl. Kontaktadressen Tab 45.1) einschließlich des Pädiatrischen Netzwerks, ein assoziiertes Mitglied (Zentrum für Klinische Studien Essen) sowie das SDGC als fester Kooperationspartner. Es kann erwartet werden, dass die Arbeit des KKSNetzwerk zukünftig auch durch die Mitarbeit der Klinischen Studienzentren bereichert wird. Für die Aufnahme weiterer Konsortialpartner ist das Netzwerk offen, wobei neue Partner an medizinische Fakultäten angebunden sein sollen.

45.2 Aufgaben der Koordinierungszentren für klinische Studien

Die Ergebnisse großer, randomisierter klinischer Studien sind eine der wichtigsten Quellen von Evidenz für die Krankenversorgung. Die Ergebnisse dieser Studien gehen zentral in die Erstellung von Leitlinien ein. Die Durchführung solcher Studien – bei großen Studienvorhaben konsortial im Netzwerk betreut – ist daher eine Hauptaufgabe der KKS. Sie betreuen in unterschiedlichem Umfang (Komplettbetreuung oder Betreuung von Teilbereichen) sowohl mono- wie auch multizentrische Studien verschiedenster Gebiete. Hierzu gehören:

- ▲ Klinische Arzneimittelprüfungen der Phasen I-IV einschließlich Arzneimittelstudien bei Kindern,
- ▲ Klinische Therapiestudien zu nicht-medikamentösen Therapieformen und Medizinprodukten sowie Therapieoptimierungs-Studien,
- ▲ Studien zur Evaluation diagnostischer Verfahren,
- ▲ Studien zur Prognose von Erkrankungen
- ▲ Studien zu gesundheitsökonomischen Fragestellungen,
- ▲ Bevölkerungsbezogene Studien zur Ermittlung von Gesundheitsrisiken.

Das KKS-Netzwerk sieht es auch als eine seiner Aufgaben an, neue Fragestellungen und Forschungsansätze (z.B. methodischer Art) in seine Arbeit zu integrieren.

2005 wurden insgesamt 426 laufende Studien im KKS-Netzwerk betreut, 117 Studien befanden sich in Vorbereitung. 80 Prozent aller laufenden Studien betrafen Arzneimittelprüfungen, 9 Prozent nicht-medikamentöse Therapien bzw. medizintechnische Produkte. Diagnostische/prognostische Studien nahmen 2 Prozent ein, epidemiologische Studien 4 Prozent und sonstige klinische Studien 5 Prozent. 76 Prozent der 2005 laufenden sowie geplanten Studien waren multizentrisch angelegt.

Tab. 45.1: Adressverzeichnis

Geschäftsstelle des KKS-Netzwerks Universität zu Köln/Klinikum c/o IMSIE, Geb. 22, Kerpener Straße 62 50937 Köln	Leitung: Insa Bruns Tel.: 0221/47 88 70 96 Fax: 0221/47 88 70 95 E-Mail: insa.brunskks-netzwerk.de Internet: www.kks-netzwerk.de
Koordinierungszentrum für Klinische Studien Berlin Charité – Universitätsmedizin Berlin Campus Virchow Klinikum Augustenburger Platz 1, 13353 Berlin	Leitung: Roswitha Bussar-Maatz Tel.: 030/450-55 30 16 Fax: 030/450 55 39 37 E-Mail: roswitha.bussar-maatzcharite.de Internet: www.kks.charite.de/
Koordinierungszentrum für Klinische Studien Dresden Technische Universität Dresden Medizinische Fakultät Carl Gustav Carus der Technischen Universität Dresden Fetscherstraße 74 01307 Dresden	Leitung: Dr. Xina Grähler Dr. med. habil. Ralph Naumann Tel.: 0351/4 58 51 60 Fax: 0351/4 58 57 99 E-Mail: xina.graehlertkksdresden.de ralph.naumannkksdresden.de Internet: www.kksdresden.de
Koordinierungszentrum für Klinische Studien Düsseldorf Medizinische Fakultät Universitätsklinikum Heinrich-Heine-Universität Düsseldorf Moorenstraße 5, 40225 Düsseldorf	Leitung: Prof. Dr. Christian Ohmann Tel.: 0211/8 11 97 01 Fax: 0211/8 11 97 02 E-Mail: christian.ohmannuni-duesseldorf.de Internet: www.kksd.de
Zentrum Klinische Studien Freiburg Universitätsklinikum Freiburg Elsässer Straße 2 79110 Freiburg	Leitung: Prof. Dr. Herbert Maier-Lenz Tel.: 0761/2 70 73 85 Fax: 0761/2 70 73 83 E-Mail: herbert.maier-lenzuniklinik-freiburg.de Internet: www.zks.uni-freiburg.de
Koordinierungszentrum für Klinische Studien Halle Martin-Luther-Universität Halle/Wittenberg Medizinische Fakultät 06079 Halle (Saale) Besucher: Technologie- und Gründerzentrum I, Haus 5, Weinbergweg 23, 06120 Halle (Saale)	Leitung: PD Dr. Dr. Jürgen Scheele Tel.: 0345/5 52 29 19 Fax: 0345/5 52 29 75 E-Mail: scheele.juergenkks-halle.de Internet: www.kks-halle.de
Koordinierungszentrum für Klinische Studien Heidelberg Ruprecht-Karls-Universität Heidelberg Universitätsklinikum Gebäude 4410, Voßstraße 2, 69115 Heidelberg	Leitung: Dr. med. Steffen Luntz Tel.: 06221/56 3 45 00 Fax: 06221/56 13 31 E-Mail: steffen.luntzmed.uni-heidelberg.de Internet: www.kks-hd.de
Koordinierungszentrum für Klinische Studien Köln Universitätsklinikum Köln Gebäude 2 a, Gleueler Straße 88 a 50931 Köln	Leitung: Dr. Ursula Paulus Tel.: 0221/4 78 79 79 Fax: 0221/4 78 79 83 E-Mail: ursula.pauluskksk.de Internet: www.kksk.de

Tab. 45.1: Fortsetzung

Koordinierungszentrum für Klinische Studien Leipzig Universität Leipzig Härtelstraße 16–18 04107 Leipzig	Leitung: Tel.: Fax: E-Mail: Internet:	Prof. Dr. Markus Löffler Dr. rer. nat. Oana Brosteanu 0341/9 71 61 250 0341/9 71 61 189 info@kksl.uni-leipzig.de www.kksl.uni-leipzig.de
Koordinierungszentrum für Klinische Studien Mainz Fachbereich Medizin der Johannes Gutenberg-Universität Mainz Langenbeckstraße 2, 55131 Mainz	Leitung: Tel.: Fax: E-Mail: Internet:	Dr. med. Monika Seibert-Grafe 06131/3 93 46 32 06131/3 93 46 33 seibert-grafe@kks-mainz.de www.kks-mainz.de
Koordinierungszentrum für Klinische Studien Marburg Philipps-Universität Marburg Robert-Koch-Straße 5 35037 Marburg	Leitung: Tel.: Fax: E-Mail: Internet:	Carmen Schade-Brittinger 06421/2 86 65 09 06421/2 86 65 17 schadebr@med.uni-marburg.de www.kks-mr.de
Koordinierungszentrum für Klinische Studien Münster Universitätsklinikum Münster Von-Esmarch-Straße 62 48129 Münster	Leitung: Tel.: Fax: E-Mail: Internet:	Dr. rer. nat. Achim Heinecke 0251/8 35 71 10 0251/8 35 70 26 kks@uni-muenster.de www.kks-ms.de
Koordinierungszentrum für Klinische Studien an den Universitätsklinikum Tübingen und Ulm/ KKS-TU GmbH (Gaststatus) Otfried-Müller-Straße 45 72076 Tübingen	Leitung: Tel.: Fax: E-Mail: Internet:	Prof. Dr. Christoph H. Gleiter 07071/2 97 22 62 07071/29 51 58 sekretariat.kks@kks-tu.de www.kks-tu.de
PAED-Net Koordinierungszentrale und PAED-Modul Mainz am KKS-Mainz Fachbereich Medizin der Johannes Gutenberg-Universität Mainz (PAED-Net) Langenbeckstraße 2, 55131 Mainz	Leitung: Tel.: Fax: E-Mail: Internet:	Dr. med. Raunhild Butzer 06131/3 93 46 46 06131/3 93 46 33 paed-net@kks-mainz.de www.paed-net.org
Assoziiertes Mitglied		
Zentrum für Klinische Studien Essen c/o IMIBE Hufelandstraße 55 45122 Essen	Leitung: Tel.: Fax: E-Mail: Internet:	Prof. Dr. Karl-Heinz Jöckel 0201/23 45 14 0201/7 23 59 33 info@zkse.de www.zkse.de
Fester Kooperationspartner		
Studienzentrum der Deutschen Gesellschaft für Chirurgie Ruprecht-Karls-Universität Heidelberg Universitätsklinikum Im Neuenheimer Feld 110, 69120 Heidelberg	Leitung: Tel.: Fax: E-Mail: Internet:	Dr. med. Christoph M. Seiler 06221/56 69 80 06221/56 69 82 sdgc@med.uni-heidelberg.de www.sdgc.de

45.2.1 Studienplanung

Schwerpunktmäßig arbeiten die KKS mit den verschiedenen Kliniken und Instituten ihrer Universitäten zusammen. Durch die darüber hinaus stattfindende Zusammenarbeit mit Partnern im außeruniversitären Bereich, wie Studienzentren und Praxisnetzwerken, kann ein breites Indikationsspektrum abgedeckt werden. Dies ermöglicht es, schnell und effektiv die krankheitsspezifische Patientenverfügbarkeit zu prüfen und eine realistische Planungsgrundlage für die jeweilige klinische Studie zu ermitteln. Der Support umfasst die Überprüfung der Durchführbarkeit (Feasibility), die Entwicklung und Konzeption von Studien, Abstimmungen mit Behörden, Übernahme von Sponsorverantwortlichkeiten im Zuge der Delegation, Projekt- und Datenmanagement, Finanzierungsplanung, die Kontaktvermittlung zu Kliniken, Forschern und Studiengruppen sowie die Auswahl und Vermittlung der Prüf- und Referenzzentren (z. B. Zentrallabor, Radiologie, Pathologie) und die Studiendokumentation.

45.2.2 Studiendurchführung

Bei der Studiendurchführung reichen die Leistungen des Netzwerks von der gezielten Hilfestellung in einzelnen Teilbereichen bis zur Komplettbetreuung von Studien. Das Leistungsspektrum umfasst allgemeine Beratung, Koordination von multizentrischen Studien, Projektmanagement inkl. Safety-Management, Qualitätssicherung durch Monitoring/Datenmanagement sowie Support durch Studienassistenten und bei Bedarf Schulung von Studienpersonal sowie die Vermittlung ausgebildeten klinischen Personals einschließlich gesonderter Räumlichkeiten.

45.2.3 Studienauswertung und Publikation

Das Netzwerk der KKS verfügt über eigene biometrische Expertise und arbeitet zudem eng mit den jeweiligen Instituten für Medizinische Statistik und Informatik der angeschlossenen Fakultäten zusammen. Dadurch wird ein hoher Standard bei der Studienauswertung, Berichterstellung und Publikation sichergestellt.

45.3 Aus- und Weiterbildung im KKS-Netzwerk

Die professionelle Betreuung von klinischen Studien durch die Zentren ist ein wichtiges Mittel, um die patientenorientierte klinische Forschung in Deutschland zu stärken, reicht jedoch auf Dauer allein nicht aus. Klinische Studien müssen, wenn sie international akzeptiert und hochrangig publiziert werden sollen, internationalen Qualitätsstandards, wie beispielsweise ICH-GCP, aber natürlich auch nationalen Gesetzen wie dem deutschen Arzneimittelgesetz (AMG) entsprechen. Um eine dauerhafte Qualitätssteigerung in der angewandten klinischen Forschung zu erzielen, sind daher zusätzliche studienbezogene Weiter- und Fortbildungsmaßnahmen erforderlich. Aus- und Weiterbildungsangebote in diesem Bereich sind ebenfalls eine zentrale Aufgabe des KKS. In den letzten Jahren wurden daher standardisierte Curricula für verschiedene Funktionsbereiche in klinischen Prüfungen entwickelt. Kurse, die nach Curriculum und Evaluation die Anforderungen des KKS erfüllen, werden mit einem entsprechenden Zertifikat gekennzeichnet, das als Qualitätssiegel für Studienassistenten, Studienleiter, Monitore und Prüfärzte fungiert.

Die Aus- und Weiterbildungsangebote des KKS sollen wissenschaftliche und organisatorische Kenntnisse und Fähigkeiten für

die Durchführung klinischer Studien vermitteln, aber auch die Motivation für die Initiierung klinischer Studien erhöhen. Darüber hinaus soll den Teilnehmern der Kurse ein Grundverständnis für EbM vermittelt werden.

Das **Curriculum „Prüfarzt“** richtet sich an Ärzte in Kliniken und Instituten, Praxen und in der Pharmaindustrie und soll die Qualifikation als Prüfarzt vermitteln. Absolventen des Kurses sollen in der Lage sein, die lokale Verantwortung für die Durchführung einer Studie in einer Klinik zu übernehmen. Dies stellt heute an Prüfarzte über die nötige medizinische Kompetenz hinausgehende spezifische Anforderungen und setzt die entsprechenden Kenntnisse auf rechtlichem, ethischem und methodischem Gebiet voraus. Die Qualifikation des Prüfarztes soll dabei nicht auf Arzneimittelstudien beschränkt sein, sondern auch andere klinische Studien mit therapeutischen, diagnostischen und prognostischen Fragestellungen umfassen.

Durch ein differenziertes Lehrangebot sollen die Anforderungen klinischer Prüfungen sowohl im niedergelassenen als auch im klinischen Bereich berücksichtigt werden.

Von 1999 bis Mitte 2006 wurden 1160 Teilnehmer in 47 Kursen weitergebildet, davon waren 45 Kurse mit 1134 Teilnehmern KKS-zertifiziert.

Das in 2006 neu hinzugekommene **Curriculum „Studienleiter“** spricht Ärzte, Zahnärzte und Naturwissenschaftler in der klinischen Forschung sowie wissenschaftliche Mitarbeiter industrieller Sponsoren an, die klinische Studien eigenverantwortlich leiten wollen. Der Kurs soll in mindestens 24 Unterrichtsstunden die Bedeutung und Einsatzmöglichkeiten aller Formen klinischer Studien als Erkenntnisinstrument in der klinischen Forschung und besondere methodische Kompetenzen in der angewandten klinischen Forschung vermitteln. Er soll die Absolventen in die Lage versetzen, in Zusammenarbeit mit methodenwissenschaftlichen

Fächern klinische Studien nach internationalen Qualitätsstandards zu konzipieren und zu leiten. Das Ausbildungsangebot zielt außer auf wissenschaftliche Kompetenz und Leitungskompetenz in einzelnen Studien auch auf die Verbesserung der Motivation für die Initiierung klinischer Studien ab.

Monitoring ist für die Qualität einer Studie im Hinblick auf Patientensicherheit und Validität der Daten unabdingbar. Klinische Monitore benötigen Erfahrung in der Vorbereitung und Durchführung von klinischen Prüfungen sowie Kenntnisse über die Inhalte der einzelnen Studien. Das ebenfalls im Jahr 2006 neu eingeführte Curriculum „Einführung in das klinische Monitoring“ vermittelt internationale Forschungsstandards und notwendige theoretische und praktische Kenntnisse und wurde auf einen Mindestumfang von 24 Unterrichtsstunden festgelegt.

Mit 120 Unterrichtsstunden nimmt das **Curriculum „Studienassistent“** den zeitlich größten Umfang der Weiterbildungsmaßnahmen des KKS-Netzwerks ein. Angesprochen wird Personal in medizinischen Assistenzberufen sowie Krankenschwestern und -pfleger (Study Nurse). Studienassistenten sollen den Prüfarzt vor Ort bei der prüfplan-gerechten Rekrutierung, Untersuchung, Behandlung und Dokumentation von Studienpatienten sowie bei der Einweisung von ärztlichem Hilfspersonal in Studienabläufe unterstützen und die lokale studienbezogene Tätigkeit in einem an der Studie teilnehmenden Zentrum koordinieren. Die Studienassistenten übernehmen Aufgaben im Rahmen mono- und/oder multizentrischer Studien vor Ort in der Institution. Dabei arbeiten sie in der Regel eng mit den Studiensekretariaten der relevanten Studien außerhalb ihrer Institution zusammen. Die Ausbildung soll für eine überwiegend selbstständige Tätigkeit qualifizieren.

Von 1999 bis Mitte 2006 wurden 1008 Teilnehmer in 40 Kursen weitergebildet,

hiervon waren 32 Kurse mit insgesamt 645 Teilnehmern KKS-zertifiziert.

Verschiedene Fachgebiete, z.B. Chirurgie, Zahnheilkunde und insbesondere Pädiatrie, stellen spezifische Anforderungen an die Studienbeteiligten. Die einzelnen KKS bieten daher auch fachbezogene Zusatzkurse und Informationsveranstaltungen an. Das KKSN beteiligt sich zudem an Masterstudiengängen (Master of Science in Medical Biometry/Biostatistics). Einige KKS haben bereits, zum Teil in Zusammenarbeit mit weiteren Partnern, Masterstudiengänge Clinical Research entwickelt. Darüber hinaus sollen Postgraduierten-Programme der Universitäten sowie neue didaktische Ansätze wie Blended-Learning und web-basiertes Training verfolgt werden. Ergänzend zu seinen Präsenzkursen wird das KKS-Netzwerk ab 2007 einen Einführungskurs zu GCP webbasiert anbieten.

45.4 Zusammenfassung und Ausblick sowie künftige Herausforderungen

Das KKS-Netzwerk hat in Deutschland einen festen Platz in der Forschungslandschaft erungen. In den letzten sieben Jahren wurden eine leistungsfähige Infrastruktur geschaffen und mehrere hundert Studienprojekte aus verschiedensten Gebieten auf vielfältigste Art wissenschaftlich wie auch organisatorisch unterstützt. Auch konnten Drittmittel für Studien aus dem BMBF/DFG-Programm erfolgreich eingeworben werden, ein Zeichen für die hohe inhaltliche wie methodische wissenschaftliche Kompetenz im Netzwerk.

Nach der degressiven Anschubfinanzierung des BMBF bis 2007 ist die Weiterführung der KKS durch Zusagen der Fakultäten sowie die Einwerbung von Drittmitteln mittel- bis langfristig gesichert. Dies betrifft

auch die Finanzierung der gemeinsamen Geschäftsstelle des Netzwerks.

Durch die Aus- und Weiterbildungsangebote des KKS-Netzwerks wird ein maßgeblicher Beitrag zur Stärkung der Studienkultur in Deutschland geleistet. Mit gezielter Nachwuchsförderung sowie der Initiierung von Masterstudiengängen soll an internationales Niveau angeknüpft werden, um attraktive Karrierewege in der klinischen Forschung zu eröffnen.

Die KKS haben mittlerweile Vorbildfunktion: An vielen Universitäten gibt es Initiativen zur Einrichtung ähnlicher Zentren. Zudem gilt das KKS-Netzwerk für viele Förderer und Studiengruppen als unparteiischer Anwalt nicht-kommerzieller Studien [3]. Durch die Kooperation mit anderen Netzwerken und Studiengruppen hat das KKS-Netzwerk seine Expertise in zahlreiche Gremien und Foren einbringen können. In Zukunft soll die Zusammenarbeit mit internationalen Organisationen – vornehmlich auf europäischer Ebene – weiter ausgebaut werden. Auch beteiligt sich das KKS-Netzwerk an dem Aufbau einer europäischen Infrastruktur zur Unterstützung multinationaler Studien.
Internet: <http://www.kks-netzwerk.de>

Literatur

- [1] Rothmund M, An international comparison of the position of clinical research in Germany. *Dtsch Med Wochenschr.* (1997), 122 (44), 1358–62
- [2] Ohmann C et al., Koordinierungszentren für Klinische Studien, Forschungsverbünde und Rahmenbedingungen für Klinische Studien in Deutschland (Kongressbericht) <http://www.egms.de/en/meetings/ecrin2004> (12.12.06)
- [3] Richter-Kuhlmann E, Eine Frage des Geldes. Nicht-kommerzielle klinische Studien. *Dtsch Arztebl.* (2006), 42, 2756–2757

46 DIMDI – Deutsches Institut für Medizinische Dokumentation und Information

Alric Rüther, Britta Göhlen

46.1 Medizininformation umfassend und unabhängig

Das Deutsche Institut für Medizinische Dokumentation und Information (DIMDI) ist eng mit der Entwicklung der Evidenzbasierten Medizin (EbM) in Deutschland verbunden. Es entwickelt effiziente Lösungen für die praxisorientierte Vernetzung von hochwertiger Medizininformation zur Unterstützung der Partner im Gesundheitswesen: Patienten und Ärzte sowie Politik, Kliniken und Krankenkassen, Forschung und Industrie. Die Unterstützung der EbM ist hier wesentlicher Bestandteil und EbM-Anwender finden beim DIMDI ein breites Spektrum von relevanten Informationen: u.a. wissenschaftliche Literatur, Daten über Arzneimittel und Medizinprodukte, Health Technology Assessment (HTA) und nicht zuletzt medizinische Begriffssysteme.

Seit seiner Gründung 1969 widmet sich das DIMDI der Vermittlung von fundiertem Medizinwissen für alle Bereiche des Gesundheitswesens. Politik und Verwaltung, Forschung und Industrie sowie Ärzte und Patienten – sie alle zählen zu den Nutzern des DIMDI. Die Informationssysteme erlauben Zugriff auf hochwertige Medizininformationen entsprechend den jeweiligen Bedürfnissen der Anwender. Interessierte können nach aktuellen Meldungen, Hintergrundinfos, Forschungsergebnissen, wissenschaftlicher Literatur oder Arbeitsgrundlagen der Medizin recherchieren. Das DIMDI hat die Evidenzbasierte Medizin in Deutschland von Beginn an begleitet, beteiligt sich intensiv an der Entwicklung und unterstützt die For-

schung in diesem Bereich. Die Ergebnisse fließen in das Informationsangebot des DIMDI ein. Für die Umsetzung einer EbM stellt das DIMDI damit eine wesentliche Informationsquelle dar. Folgende Informationssysteme stehen zur Verfügung:

- ▲ Informationssystem Gesundheitswesen,
- ▲ Informationssystem Health Technology Assessment (HTA),
- ▲ Informationssystem Arzneimittel,
- ▲ Informationssystem Medizinprodukte,
- ▲ Medizinische Klassifikationen und Begriffssysteme.

Das DIMDI ist eine nachgeordnete Behörde des Bundesministeriums für Gesundheit (BMG). Das Spektrum der Aufgaben hat sich im Laufe der Jahre wesentlich vergrößert und an Bedeutung gewonnen. Heute ist das Institut in wichtige Gesetze und Verordnungen der Gesundheitspolitik wie auch in die Politikberatung eingebunden. Als moderner Informationsanbieter und IT-Plattform koordiniert das DIMDI auch IT-Projekte im Geschäftsbereich des BMG.

46.2 Informationssystem Gesundheitswesen

Den Kern des Informationsangebots bilden rund 70 medizinische Datenbanken mit 100 Mio. Nachweisen und Zusammenfassungen wissenschaftlicher Literatur oder Fakten aus der Medizin und ihren Randbereichen (s. Abb. 46.1). Die Hälfte der Datenbanken ist kostenfrei zu recherchieren. Neben großen internationalen Datenbanken wie MED-

Abb. 46.1: Das Informationsangebot des DIMDI: www.dimdi.de

LINE, EMBASE und SciSearch und vielen speziellen Datenbanken wie der Cochrane Library oder Bewertungen von Alternativen zu Tierversuchen (AnimAlt-ZEBET) bietet das DIMDI exklusiven Zugriff auf Daten zur Zulassung von Arzneimitteln (AMIS) und Anzeige von Medizinprodukten. Aktuelle Meldungen und die amtlichen Teile des Bundesanzeigers mit Informationen aus dem Ressort des BMG sind recherchierbar. Außerdem ermöglicht das DIMDI die Recherche in den Artikeln des Deutschen Ärzteblatts und im multimedialen Angebot der Virtuellen Bibliothek für Medizin.

EbM: Zuverlässig recherchieren

Die Evidenzbasierte Medizin (EbM) benötigt aktuelle Information zum Zeitpunkt der Entscheidung, beispielsweise im Rahmen der Be-

ratung für einen Patienten. Für diesen Zweck stellt das DIMDI eine Palette mächtiger Werkzeuge bereit, die eine schnelle und effektive Suche im großen Angebot an wissenschaftlichen Quellen ermöglichen. Für die Datenbankrecherche bietet das DIMDI verschiedene Zugänge an: Die menügeführte Oberfläche DIMDI-SmartSearch bietet einen komfortablen Zugriff auf die Datenbanken, die Kommandosprache DIMDI-ClassicSearch wird überwiegend von Recherche-Professionals für sehr komplexe Fragestellungen genutzt.

So genannte Superbases, Datenbanken-Gruppen zu verschiedenen Themenbereichen, erlauben die zuverlässige Recherche nach spezifischen Themen, z.B. für die Fachrichtungen Medizin, Arzneimittel oder HTA (s. Tab. 46.1). Das Retrievalsystem des DIMDI sucht automatisch über alle Datenbanken der jeweiligen Gruppe und filtert Duplikate

Tab. 46.1: Einige Datenbanken zum Thema HTA, die sich beim DIMDI gleichzeitig recherchieren lassen

Cochrane Library – CDMR (CDMR)	MEDLINE ME60
Cochrane Library – CDSR (CDSR93)	DARE (CDAR94)
Cochrane Library – Central (CCTR93)	HTA (INAHTA)
DAHTA-Datenbank (DAHTA)	NHS-EED (NHSEED)
EMBASE (EM74)	SciSearch (IS74)

(preisgünstigere Fundstellen bevorzugt; Premiumrecherche). Damit entfällt die umständliche Suche über mehrere Datenbanken mit zeitaufwendigem Vergleich der Treffer. Relevante Artikel können im Weiteren online über angeschlossene Bibliotheken als Volltext bestellt werden oder sind bereits über das Volltextarchiv abrufbar, das das DIMDI in Kooperation mit der Deutschen Zentralbibliothek für Medizin (ZB MED) pflegt. Darin nicht enthaltene Volltexte sind über einen Linkresolver zu anderen Fundstellen (z.B. bei Verlagen) direkt verlinkt.

In Recherche- und Datenbankkursen können Einsteiger den Umgang mit den Suchsystemen und speziellen Datenbanken erlernen. Ob für die schnelle Suche während der Visite auf der Station oder das ausgefeilte Retrieval eines systematischen Reviews: Jeder kann seinen Bedürfnissen entsprechend das Informationssystem nutzen.

46.3 Informationssystem Health Technology Assessment: Wissenschaft für Politik

Health Technology Assessment (HTA) gründet auf dem Konzept der EbM. Auch hier geht es darum, den aktuellen Stand der Wissenschaft so zur Verfügung zu stellen, dass er zeitnah als Grundlage für Entscheidungen genutzt werden kann. Im Fokus steht jedoch nicht der individuelle Fall (wie bei der EbM), sondern die gesundheitliche Versorgung der Bevölkerung. HTA dient der wissenschaftlichen Unterstützung gesundheitspolitischer Fragestellungen. Damit ist ein anderer Infor-

mationsumfang gefordert. Neben der medizinischen Effektivität unter optimalen Bedingungen, wie sie die wissenschaftlichen Studien zeigen, sind eine Bewertung der Effektivität eines Verfahrens unter Alltagsbedingungen, die Betrachtung der gesundheitsökonomischen Effizienz und die Beurteilung ethischer oder sozialer Implikationen erforderlich. In HTA-Berichten sind diese Arbeiten synthetisiert und aufbereitet.

HTA-Berichte sind von hoher Qualität, entsprechend systematischen Reviews. Sie lassen sich sehr gut als Grundlage für Entscheidungen in der EbM heranziehen. Die Untersuchungen zur Wirksamkeit sind hier sicher von primärem Interesse. Die Informationen zur Gesundheitsökonomie und Ethik sind in vielen Fällen auch bei der individuellen Patientenberatung hilfreich. Zudem gibt es bei den meisten HTA-Berichten ausführliche Zusammenfassungen, so genannte Executive Summaries, die das Ergebnis und die Qualität eines Berichtes verständlich erfassen – als schnellen Einstieg für die Gesundheitspolitik. Die Executive Summaries geben kurz und prägnant qualitativ hochwertige Information über ein medizinisches Verfahren. Auch für die EbM sind sie eine schnelle und verlässliche Informationsquelle.

In der deutschen Gesundheitspolitik spielt HTA seit Mitte der 90er-Jahre eine zunehmende Rolle. 2000 wurde die Deutsche Agentur für HTA des DIMDI, DAHTA@DIMDI, gegründet. Sie betreibt ein Informationssystem HTA und führt ein Programm zur Erstellung von HTA-Berichten durch. DAHTA ist in die internationalen HTA-Programme integriert und arbeitet entspre-

chend den internationalen methodischen Standards. Sie ist Mitglied im International Network of Agencies for HTA (INAHTA), des Netzwerks der öffentlichen HTA-Organisationen. Außerdem ist DAHTA aktiv in der internationalen Fachgesellschaft HTA international (HTAi), der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) und dem Deutschen Netzwerk Evidenzbasierter Medizin (DNEbM). DAHTA ist darüber hinaus einer der führenden Partner bei Aufbau und Betrieb eines Europäischen Netzwerks für HTA (EUnetHTA).

Im Rahmen des HTA-Programms entstehen beim DIMDI HTA-Berichte nach einem qualitätsgesicherten standardisierten Verfahren. Die Berichte treffen Aussagen zu Nutzen, Risiko, Kosten sowie Auswirkungen medizinischer Technologien mit Bezug zur gesundheitlichen Versorgung der Bevölkerung. Dabei fallen unter den Begriff Technologie sowohl Medikamente als auch Instrumente, Geräte, Prozeduren, Verfahren sowie Organisationsstrukturen. Vorrang haben Themen, für die gesundheitspolitischer Entscheidungsbedarf besteht. Zur Zielgruppe gehören sämtliche Entscheidungsgremien im deutschen Gesundheitswesen, aus Gesundheitspolitik, Krankenhausmanagement, medizinischen Fachberufen oder dem Patienten-umfeld.

Basis der Themenfindung der HTA-Berichte ist die öffentlich zugängliche Themensammlung, in die jeder eine gewünschte Frage mit Hintergrundinformationen über einen Fragebogen vorschlagen kann (www.dimdi.de – Link HTA). Die Themen werden zwei Mal im Jahr ausgewertet. Das HTA-Kuratorium legt schließlich die Themen und Prioritäten für die zu erstellenden Arbeiten in einem mehrstufigen Verfahren fest. Die Mitglieder des HTA-Kuratoriums bilden einen repräsentativen Querschnitt aus den Gremien des deutschen Gesundheitssystems (die Selbstverwaltung, d.h. Krankenkassen, Krankenhäuser und Ärzte), ergänzt durch

Verbraucher- oder Patientenvertreter sowie Repräsentanten des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) und aus der Industrie als Beobachter.

Das Verfahren der Entstehung der HTA-Berichte wird kontinuierlich geprüft und weiterentwickelt, unter Einbezug internationaler Standards. Die Berichte werden publiziert über das elektronische Journal GMS-HTA (www.egms.de) sowie über die DAHTA-Datenbank (www.dimdi.de). Sie sind kostenfrei im Volltext über diese Quellen erhältlich. Das GMS-HTA-Journal bietet zudem ergänzende Fachartikel aus dem Bereich HTA.

Das HTA-Programm ist Teil des Informationssystems HTA, das kostenfrei aktuelle medizinische Informationen für Fachkreise und Laien bietet. Den Kern des Angebots bilden die als Volltext abrufbaren deutschsprachigen HTA-Berichte, die im Auftrag von DAHTA@DIMDI herausgegeben werden. Weiterhin sind nationale HTA-Berichte anderer Herausgeber sowie internationale HTA-Berichte im Original mit deutscher Übersetzung der Zusammenfassungen verfügbar, z.B. des britischen National Institute for Clinical Excellence (NICE). Die DAHTA-Datenbank mit den HTA-Berichten ist auch in das Informationssystem Gesundheitswesen integriert, sodass sich die Berichte gleichzeitig mit anderen Literaturdatenbanken des DIMDI recherchieren lassen.

Seit 2000 veranstaltet DAHTA@DIMDI jährlich ein HTA-Symposium zu aktuellen Fragestellungen und bietet damit eine Austauschplattform für Experten und Anwender. Im Rahmen des Symposiums spielt auch die Fortbildung eine wichtige Rolle. In den Jahren 2004 und 2005 kamen über 250 Teilnehmer zu den Symposien, die durch die Ärztekammer Nordrhein als ärztliche Fortbildung zertifiziert sind. Aktuelle Informationen über Fokus und Zeitpunkt des nächsten Symposiums sind über die DIMDI-Webseite und den HTA-Newsletter erhältlich.

46.4 Informationssystem Arzneimittel

Als National Competent Authority ist das DIMDI auf europäischer Ebene der nationale Ansprechpartner auf dem pharmazeutisch-telemedizinischen Gebiet für Informationsübermittlung, Informationsdarstellung und Dokumentation im Arzneimittelbereich. Auf Grundlage des § 67a des Arzneimittelgesetzes (AMG) hat das DIMDI den Auftrag, ein zentrales datenbankgestütztes Informationssystem für Arzneimittel zu errichten. Zusätzlich zu amtlichen Zulassungsinformationen aus dem Arzneimittel-Informationssystem der bundesdeutschen Zulassungsbehörden (AMIS) gibt es Datenbanken mit umfassenden Informationen zu Anwendung, Neben- und Wechselwirkungen von Arzneistoffen und Fertigarzneimitteln.

Ein Meilenstein auf dem Weg zur Vernetzung der Arzneimittelinformationen ist die spezielle Anwendung DIMDI PharmSearch. Auf Grundlage von Arzneimitteldaten des ABDAMED-Datenbestands können Wirkstoffe, Handelsnamen, Zusammensetzung, Dosierung, Indikationen, Kontraindikationen oder Nebenwirkungen von Medikamenten aufgerufen werden. Die DIMDI PharmSearch ist auf die Fragestellungen von Fachkreisen ausgerichtet, die kostenfrei Zugriff haben (DocCheck-Passwort erforderlich). Über direkte Verlinkung sind aus dem jeweiligen Dokument zum Fertigarzneimittel zusätzliche Informationen aus der Roten Liste und der Gelben Liste Identia (u.a. Abbildungen von Darreichungsformen) abrufbar. Der monatlich aktualisierte Datenbestand umfasst alle apothekenüblichen Arzneimittel in Deutschland, also auch Phytopharmaka und Homöopathika. Dies entspricht weit über 100 000 Arzneipackungen. Ein Interaktionscheck überprüft automatisch mögliche Wechselwirkungen zwischen verschiedenen Arzneimitteln. Im Rahmen einer EbM-basierter Patientenbetreuung sind diese Informa-

tionen außerordentlich wertvoll in Ergänzung zu den Ergebnissen der wissenschaftlichen Studien.

46.5 Informationssystem Medizinprodukte

Das Informationssystem Medizinprodukte des DIMDI ist europaweit das erste vollelektronische Anzeigeverfahren für Medizinprodukte mit automatischen Benachrichtigungssystemen. Es enthält unter anderem Daten zu Medizinprodukten, In-vitro-Diagnostika, Klinischen Prüfungen/Leistungsbeurteilungen, Bescheinigungen (Zertifizierungen) und Vorkommnissen mit Medizinprodukten. Das Informationssystem besteht aus drei Komponenten:

- ▲ Erfassungssystemen, mit denen Hersteller, Bevollmächtigte, Prüfeinrichtungen oder Prüfstellen (Benannte Stellen) die Anzeigen eingeben und bearbeiten.
- ▲ Datenbanken mit den regulatorischen Daten jeweils für Medizinprodukte und In-vitro-Diagnostika sowie Adressdatenbanken mit den Adressen der Anzeigenden, der zuständigen Behörden und der benannten Stellen.
- ▲ Benachrichtigungssystemen für den Informationsfluss zwischen Anzeigenden und zuständigen Behörden.

Ein Teil der Daten steht als „Medizinprodukte – Öffentlicher Teil“ über die Datenbankrecherche des DIMDI jedem Interessierten zur Verfügung. Insgesamt enthalten acht weitere Datenbanken Informationen zu dieser Thematik. Im Sinne der EbM lassen sich hier bei Bedarf Hintergrundinformationen zu Medizinprodukten recherchieren. Sie können für die Bewertung der Therapieoptionen insbesondere im Hinblick auf die individuellen Bedürfnisse und Anforderungen des Patienten genutzt werden.

46.6 Medizinische Klassifikationen und Begriffssysteme

Neben den für EbM herausragenden Informationssystemen im Bereich medizinisch-wissenschaftlicher Literatur betreut und entwickelt das DIMDI Begriffssysteme, die eine EbM indirekt unterstützen: ICO, ICF, ICD-O, OPS, ATC/DDD, UMDNS etc. Mit den medizinischen Klassifikationen liefert das DIMDI beispielsweise die Grundlagen für neue Ab-

rechnungssysteme wie die Diagnosis Related Groups. Als WHO Collaborating Centre ist das DIMDI an der Weiterentwicklung der ICD (International Classification of Diseases) auf internationaler Ebene beteiligt. Das Institut betreut darüber hinaus Begriffs- und Ordnungssysteme wie Alpha-ID, LOINC und OID, die wichtig für Telematikanwendungen im Gesundheitswesen sind.

Internet: <http://www.dimdi.de>

VII Ressourcen

47	Bias, Confounding, Chance	483
	<i>Jürgen Windeler</i>	
48	Checklisten zur Bewertung von klinischen Studien und Leitlinien	491
	<i>Guido Lerzynski, Hilke Bertelsmann, Regina Kunz</i>	
49	Glossar zur Evidenzbasierten Medizin	497
	<i>Regina Kunz, Dagmar Lühmann, Jürgen Windeler, Monika Lelgemann, Norbert Donner-Banzhoff</i>	

47 Bias, Confounding, Chance

Jürgen Windeler

Die Praxis Evidenzbasierter Medizin beruht darauf, sich systematisch einen fundierten Einblick in die internationale Literatur zu einer konkreten Problemstellung zu verschaffen. Für die Bewertung der Quellen lernt man die einschlägigen Qualitätskriterien für therapeutische, diagnostische und prognostische Studien kennen und sammelt Erfahrungen mit den zahlreichen publizierten Checklisten und den verschiedenen Evidenzklassierungen.

Das Ziel aller Bemühungen, Studien zu bewerten und zu beurteilen, ist es, zu entscheiden, ob ein mitgeteiltes Ergebnis als **valide** angesehen werden kann oder nicht.

47.1 Interne und externe Validität

Bezüglich der Validität von Studien muss man zwei Komponenten unterscheiden: die **interne** und die **externe** Validität. Die Unterscheidung ist von sehr großer Bedeutung, auch wenn sie in der Praxis häufig vernachlässigt wird, was zu Missverständnissen und Missinterpretationen führt. Dazu zunächst ein konkretes Studienbeispiel:

In der Studie „Morbidity and Mortality after Stroke Eprosartan compared with Nitrendipine for secondary prevention (MOSES)“ [1] wurden Patienten mit einem vorausgegangenen cerebrovaskulären Ereignis (Schlaganfall/TIA), die außerdem eine behandlungsbedürftige arterielle Hypertonie aufwiesen, in zwei Gruppen randomisiert. Die eine Gruppe erhielt anfangs Eprosartan in einer initialen Dosis von 600 mg pro Tag, die Vergleichsgruppe eine initiale Behand-

lung von 10 mg Nitrendipin pro Tag. Die beiden Medikationen konnten gesteigert und bei unzureichender antihypertensiver Wirkung um weitere Antihypertensiva ergänzt werden. Die Studie wurde offen, aber mit einer verblindeten Bewertung der Endpunkte durchgeführt. Die Patienten wurden über mindestens 2 Jahre beobachtet. Von den 1405 randomisierten Patienten zogen 53 (29/24) ihr Einverständnis vor der ersten Medikamenteneinnahme zurück. Von den restlichen 681 Patienten in der Eprosartan-Gruppe hatten 80 ein cerebrovaskuläres Ereignis im Verlauf der Studie. Unter den 671 Patienten der Nitrendipin-Gruppe waren es 89 Studienteilnehmer. (Der eigentliche primäre Endpunkt war eine Kombination aus verschiedenen, auch repetitiven Ereignissen von TIA bis Tod.)

Aus den Angaben zu den cerebrovaskulären Ereignissen lässt sich ein Relatives Risiko von 0,88 mit einem 95%-Konfidenzintervall von 0,65 bis 1,2 berechnen.

Im Sinne der internen Validität ist bei einer solchen Studie die Frage zu klären: Sind die Ergebnisse zu der **Frage, die sich die Studie gestellt hat**, als valide anzusehen? Valide bedeutet dabei eine adäquate Kontrolle möglicher Confounder (s.u.), die Minimierung von Bias und die Kontrolle (im Sinne einer Quantifizierung) des Zufalls, genauer: der zufälligen Fehlermöglichkeiten. In der genannten Studie bedeutet dies: Ist der berechnete Unterschied zwischen den beiden Therapien als valide Aussage unter den hier gewählten Studienbedingungen (Indikationen, Wahl der Kontrollgruppe, Dosis, klinische Endpunkte usw.) anzusehen?

Eine davon sorgfältig zu unterscheidende Frage ist, ob die Fragestellung einer Studie überhaupt sinnvoll oder relevant ist. Dies kann z.B. die gewählte Kontrollbehandlung, deren Dosis oder die Einschlusskriterien betreffen. Diese Bewertung ist keine Frage der internen, sondern der externen Validität. Sie betrifft also die Abschätzung, inwieweit die Ergebnisse der Studie **über die in der Studie abgebildete Situation hinaus** als relevant und verwendbar anzusehen sind.

Als (Extrem-)Beispiel können sorgfältig geplante und durchgeführte, methodisch aufwendige Tierversuche gelten, die eine sehr hohe interne Validität haben, indem sie also eine Aussage über die untersuchte Tierpopulation unter bestimmten Versuchsbedingungen machen. Ihre externe Validität für die medizinische Praxis wird üblicherweise sehr gering sein.

Die Unterscheidung ist deshalb von großer Bedeutung, weil die bekannten methodischen Instrumente, die in den Checklisten und Evidenzklassierungen abgebildet werden, nur der Sicherung der internen Validität von Studien dienen. Die externe Validität ist von dem mit diesen Instrumenten angestrebten Schutz vor verzerrten Aussagen (Bias) grundsätzlich nicht betroffen. Methodischen Schutz im engeren Sinne gibt es in Bezug auf die externe Validität in der medizinischen Forschung, speziell der Anwendungsforschung, nicht.

Bei der Beurteilung der externen Validität handelt es sich also um ein fachliches Ermessen mit mehr oder weniger scharfen Kriterien. In der o.g. Studie ist etwa die Frage zu stellen, ob die Vergleichstherapie Nitrendipin mit der verwendeten Dosis eine breit (oder überhaupt) angewandte Therapie ist, sodass also der Vergleich für die Praxis eine Bedeutung hätte.

Es gibt in der Evidenzbasierten Medizin eine häufige Diskussion darüber, ob bei der Bewertung von Studien der internen Validität oder der externen Validität eine größere

Bedeutung beizumessen ist. Obwohl eine Entscheidung darüber bei der einzelnen Studie nicht immer leicht zu fällen ist, so gibt es ein entscheidendes Gegenargument gegen eine besondere Wertschätzung der externen Validität im Sinne einer „Übertragbarkeit“ von Studienergebnissen, auch zulasten der internen Validität. Dieses Argument lautet in Kurzform: ohne interne Validität keine externe Validität. Eine Studie, die ein hohes Risiko birgt, verzerrte Aussagen zu machen, mag zwar zu einer Anwendungssituation theoretisch passen, aber für diese dennoch keine Aussage zulassen, da mit möglicherweise grob verzerrten Ergebnissen grundsätzlich keine **Entscheidungsunterstützung** – ganz gleich, wie gut die Situation auch übertragbar sein mag – möglich ist.

47.2 Bias

Die Bewertung der internen Validität lässt sich in der Frage zusammenfassen: Inwieweit ist in der vorliegenden Studie ein Bias ausgeschlossen bzw. minimiert? Bias bedeutet dabei eine Verzerrung, eine Verfälschung eines Ergebnisses. Diese Verzerrung resultiert daraus, dass neben den Faktoren, deren Wirkung man untersuchen will, andere Faktoren auf das Ergebnis wirken, sodass eine Beschreibung des Effektes des interessierenden Faktors (typischerweise einer Behandlung) durch das Wirken der anderen Faktoren gestört wird. Diese störenden Faktoren bezeichnet man als Confounder (Faktoren, die „Verwirrung“ stiften), den Prozess als Confounding. Ziel aller Bemühungen zur Sicherung der internen Validität ist es also, ein Confounding möglichst auszuschalten. Das bedeutet natürlich nicht, die Confounder selbst abzuschaffen (z.B. das Alter, das Geschlecht, das Krankheitsstadium, die Krankheitsdauer und vieles andere), sondern nur deren (differenzielles) Wirken auf das Ergebnis zu unterbinden.

- Die bekannten Maßnahmen
- ▲ zufällige Zuteilung von Patienten zu Therapiegruppen (Randomisierung),
 - ▲ ein die Randomisierung ergänzendes Concealment (verdeckte Zuordnung der Patienten zur Interventions- oder Kontrollgruppe),
 - ▲ Verblindung von möglichst vielen beteiligten Parteien (Einfach-, Doppel-, Dreifachblindheit),
 - ▲ Einbinden aller in eine Studie eingeschlossenen Patienten in die Auswertung (Intention to treat),
- sind die in Therapiestudien wichtigsten methodischen Instrumente, um Bias zu vermeiden, also die interne Validität einer Studie sicherzustellen. Zur Wiederholung: Der Begriff „Bias“ ist nur im Zusammenhang mit der Bewertung der internen Validität sinnvoll, nicht für die externe Validität.

47.3 Prinzip statistischer Aussagen

Methodisch wird zwischen zwei zwar verwandten, aber im Prinzip zu unterscheidenden Zielsetzungen bei Aussagen aus Studien unterschieden: dem **Testen** und dem **Schätzen**. Beim Testen interessiert man sich nur für die grobe Frage: Ist A anders (besser) als B? Dies stellt die Grundfrage aller klinischen Studien zu Wirksamkeit und Schaden dar. Die Bearbeitung erfolgt mit so genannten Signifikanztests und die – legendäre – Antwort auf die Frage besteht in der Aussage „Ja“ (denn das Ergebnis ist „signifikant“) – oder „Nein“ (denn das Ergebnis ist „nicht signifikant“).

Trotz der Allgegenwärtigkeit von „Signifikanzen“ ist genau genommen aber das **Schätzen** die wichtigere Zielsetzung. Dies hat zum einen den Grund, dass Schätzen und Testen äquivalent in dem Sinne sind, dass sie in den meisten Fragestellungen zum gleichen Ergebnis führen, das Schätzen aber den großen Vorteil hat, ergänzend zu der

groben Ja/Nein-Antwort des Testens auch noch eine quantitative Aussage zur Größe eines Unterschiedes zu liefern – umgekehrt gilt dies nicht. Das Schätzen (mithilfe von Konfidenz-Intervallen) ist zum anderen deshalb vorzuziehen, da bei der Bewertung von therapeutischen und diagnostischen Verfahren eben genau die Frage des Ausmaßes eine wichtige Rolle spielt, wie z.B. in den umfangreichen Diskussionen über die Kommunikation von Ergebnissen mit relativen oder absoluten Effektmaßen deutlich wird. Bei diagnostischen Tests, bei denen es auf die möglichst zuverlässige Bestimmung z.B. von Sensitivität und Spezifität ankommt, ist sogar das Schätzen von ganz zentraler, das Testen von völlig nachrangiger Bedeutung.

Der Begriff „Schätzen“ ist in diesem Zusammenhang nicht im umgangssprachlichen Wortsinne zu verstehen. Es geht hier nicht um ein Vorgehen „über den Daumen“, sondern um ein rechnerisches Verfahren mit einer klaren Zielsetzung. Um das Prinzip zu illustrieren, soll eine Situation betrachtet werden, in der mit Pfeil und Bogen auf eine runde Zielscheibe geschossen wird. Aufgabe beim Schießen ist es, genau die Mitte zu treffen. Die Abbildung 47.1a–c illustriert zunächst den Unterschied zwischen einer zufälligen Streuung und einem verzerrten Ergebnis.

In Abbildung 47.1a liegen – trotz sorgfältigen Zielens – alle Ergebnisse nicht **auf** der Mitte, sie streuen in ihrer Umgebung. Die Ergebnisse liegen jedoch gleichmäßig um die Mitte herum (würde man einen Mittelwert über die Ergebnisse bilden, so würde dieser dicht an oder auf der Mitte liegen). Die Abweichungen der Einzelwerte von der Mitte entstehen nur durch zufällige Schwankungen.

In Abbildung 47.1b ist ein Trefferergebnis mit gleicher Streuung dargestellt, dessen Mitte aber nicht in der Nähe des Mittelpunkts der Zielscheibe liegt. Würde man für diese Ergebnisse einen Mittelwert berech-

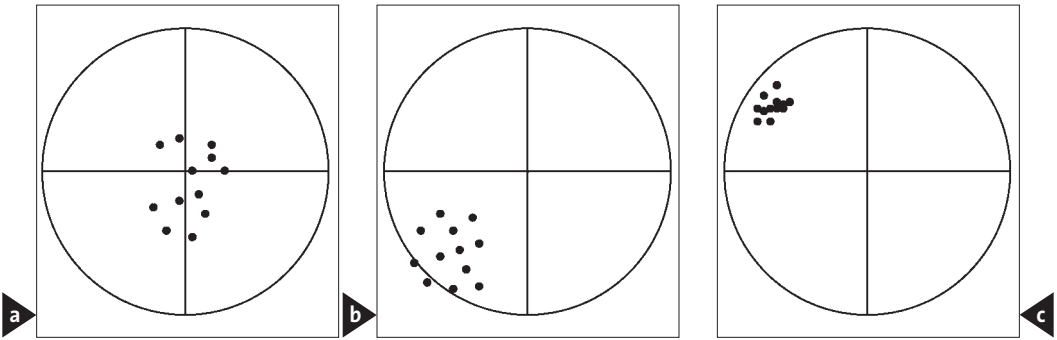


Abb. 47.1a–c: Unterschied zwischen einer zufälligen Streuung und einem verzerrten Ergebnis

nen, so läge er – trotz Zielens auf die Mitte! – deutlich neben dem Zentrum der Zielscheibe. Es ergäbe sich also bei den Ergebnissen neben den zufälligen Schwankungen zusätzlich eine systematische Abweichung (Verzerrung, Bias) vom Ziel. Vorstellbar ist, dass hier ein stetiger Seitenwind gewirkt hat, die Visiereinrichtung des Bogens verstellt war, dass der Schütze einen Sehfehler hat oder abgelenkt war.

Abbildung 47.1c zeigt, dass auch Ergebnisse mit geringer Streuung verzerrt sein können. Deren Mitte weicht also ebenfalls systematisch von der Mitte der Zielscheibe ab.

Damit nun deutlich wird, inwiefern die Bewertung von zufälligen und systematischen Abweichungen von Ergebnissen insbesondere bei klinischen Studien relevant ist, ersetzen wir die Zielscheibe durch eine große weiße Holzwand, auf der irgendwo eine Fliege sitzt. Es soll versucht werden, mit dem Pfeil die Fliege zu treffen. Das Problem: In dem Moment, wo sich der Pfeil der Scheibe nähert, fliegt die Fliege weg. Die Aufgabe ist nun nicht mehr, die Kongruenz zwischen Ziel (das unbekannt ist) und Ergebnissen festzustellen, sondern – die typische Situation klinischer Studien und empirischer Forschung generell – aus den Ergebnissen zu schließen, wo das Ziel war. Stellen wir uns vor, dass wir die Fliege nach einem Schuss immer wieder an die gleiche Stelle der Holzwand als Ziel für folgende Schüsse setzen

könnten, dann könnten wir auch eine größere Zahl von Schüssen und deren Ergebnis als Anhaltspunkt, als nach definierten Regeln gewonnene „Schätzung“ für den Fliegensitzplatz verwenden.

In diesem Szenario wird deutlich, dass die Möglichkeiten – und damit die Herausforderung – für das Treffen einer validen Aussage („Hier ist der Fliegensitzplatz.“) nur in der sorgfältigen Definition der Rahmenbedingungen des Bogenschießens liegen. Die Möglichkeit, eine Verzerrung im Ergebnis zu überprüfen, besteht nicht, da keine Referenz mehr existiert. Die Fliege ist weg.

Man wird also den Schützen möglicherweise bzgl. seiner Sehfähigkeit untersuchen, dafür sorgen, dass er nicht abgelenkt wird, den Bogen und die Pfeile sorgfältig prüfen, nur bei Windstille schießen und auch sonst alle störenden Bedingungen, die die Treffer systematisch ablenken könnten, zu beseitigen suchen. Man wird sich bemühen, alle **potenziellen Confounder** auszuschließen oder, wenn dies nicht möglich ist, ihre Auswirkungen unter Kontrolle zu bringen. So könnte man bei einem stetigen Seitenwind einen Vorhalt einplanen, um damit den Einfluss des Windes zu kompensieren.

47.4 Den Zufall im Griff

Neben dem essenziellen Ziel einer Studie, durch geeignete Maßnahmen zur Vermei-

dung von verzerrten Ergebnissen (Bias) zu validen Aussagen zu kommen, dient eine Studie auch dazu, den Einfluss des Zufalls auf die Ergebnisse in Grenzen zu halten. Bemühen wir noch einmal die Fliege: Wenn man unter sorgfältigster Beachtung aller Maßgaben zur Sicherung der internen Validität einmal auf die weiße Wand schießt, dann wird man sagen (oder eher: hoffen) können, dass die Fliege „in der Nähe“ des Treffers gegessen hat. Wie groß diese „Nähe“ aber ist, weiß man nicht. Schießt man jedoch mehrmals auf die Scheibe (immer unter „Ideal“-Bedingungen zur Vermeidung von Bias), so kann man schon eher den Bereich eingrenzen, wo die Fliege gegessen haben könnte; je mehr Schüsse man verwendet, desto genauer. Dabei mag auch deutlich werden, dass eine große Zahl von Schüssen (oder Patienten in Studien) allein nichts über die Richtigkeit von Ergebnissen aussagt. Die große Zahl kontrolliert nur die zufällige Streuung, nicht die systematische Abweichung.

Die statistische Formel, um aus den Treffern auf einer Scheibe auf die Position der Fliege zu schließen, um also den Fliegensitzplatz zu „schätzen“, ist das so genannte Konfidenzintervall (auch Vertrauensintervall oder Vertrauensbereich genannt). Es beschreibt, abhängig von der Zahl der Versuche und der „natürlichen“ Streuung der Ergebnisse (ein Anfänger wird eine höhere natürliche Streuung haben als ein Profi und dieser eine höhere Streuung als eine Maschine), in welchem Bereich die Fliege wahrscheinlich gegessen hat. Die Formel für ein einfaches 95%-Konfidenzintervall (für einen Mittelwert x) lautet:

$$x \pm 1.96 \frac{s}{\sqrt{n}}$$

x = Mittelwert, s = Standardabweichung, n = Fallzahl/Stichprobenumfang; der Wert 1.96 ergibt sich für ein 95%-Konfidenzintervall (in bestimmten Situationen ist hier ein anderer Wert erforderlich, für die Praxis kann man sich aber „über den Daumen“ den Wert 2 merken); für ein 99%-Konfidenzintervall beträgt der Wert 2.6, für ein 90%-Konfidenzintervall 1.64 (es handelt sich dabei um Quantilswerte der Standardnormalverteilung).

Die Formel führt zu zwei Zahlen als Begrenzung eines Intervalls (auf einer eindimensionalen Skala). Dieses Intervall enthält mit 95% Wahrscheinlichkeit (also leider nicht immer!) den Ort, an dem die Fliege in diesem Versuch gegessen hat. Anders ausgedrückt: Wenn man sehr viele solcher Versuche macht und jeweils ein 95%-Konfidenzintervall berechnet, so liegt in 95% aller dieser Versuche der tatsächliche Fliegensitzplatz innerhalb des Intervalls. Wie man der Formel des Konfidenzintervalls ansieht, wird es

- ▲ mit steigender Streuung breiter (s steht im Zähler), die Aussage wird also ungenauer,
- ▲ mit steigender Versuchszahl schmaler, also genauer (n steht im Nenner),
- ▲ mit erhöhtem Sicherheitsbedürfnis breiter (ausgedrückt in der Konfidenzwahrscheinlichkeit).

Ein solches Konfidenzintervall (üblich ist in Analogie zu einem Signifikanzniveau von 5% ein 95%-Konfidenzintervall) lässt sich für jede in der EbM gebräuchliche Maßzahl berechnen. Häufig kann man dies als Leser aus den in den Studienpublikationen oder -berichten angegebenen Zahlen auch selbst tun. Die Berechnung für Mittelwerte, Differenzen von Mittelwerten, Relative Risiken, Odds-Ratios und NNT ist verhältnismäßig einfach. Entsprechende Formeln findet man in einschlägigen Lehrbüchern. Konfidenzintervalle für einzelne Raten, dies betrifft Erfolgsraten, Risiken oder die Testkriterien bei diagnostischen Tests, Sensitivität und Spezifität, sind insbesondere bei kleinen Fallzahlen schwieriger zu berechnen, weil eine approximative Formel (die auf der Normalverteilung basiert) hier nicht zuverlässig anwendbar ist. Auch viele Standardprogramme sind nicht ausreichend. Eine gute Möglichkeit bietet <http://statpages.org/confint.html> bzw. die altbewährten Geigy-Tabellen.

Konfidenzintervalle haben in der medizinischen Forschung, in der Mitteilung und Bewertung von Studienergebnissen und da-

mit in der Praxis Evidenzbasierter Medizin eine ganz zentrale Bedeutung, wie die folgende Tabelle illustriert. Die Angabe der Sensitivität eines diagnostischen Tests mit 80% ist wenig nützlich, wenn keine Informationen über die Validität dieser Zahl mitgeteilt werden. Dabei ist einerseits die Abwesenheit von Verzerrungsmöglichkeiten zu bewerten, zum anderen aber die Zuverlässigkeit der Zahl an sich: Können wir die Aussage treffen, dass die Fliege bei 80, vielleicht auch 79 oder 81 gesessen hat oder könnte es auch 46 oder gar 23 gewesen sein? Für diese Aussage ist das Konfidenzintervall unerlässlich, das die Genauigkeit der Angabe (des **Schätzers** oder **Schätzwerts**) 80% quantifiziert und so eine klinische Bewertung überhaupt erst möglich macht. Die Konfidenzintervalle in Tab. 47.1 sollten sehr zu denken geben: Die meisten Diagnosestudien sind viel zu klein!

Auch bei dem Vergleich von Behandlungen ist eine solche Quantifizierung hilfreich,

wenn nicht gar notwendig: Die Aussage, unter der Behandlung überlebten doppelt so viele Patienten wie in der Kontrollgruppe (**Schätzer** für das relative Risiko 0.5) wird erst dadurch beurteilbar, dass (Vermeidung von Bias vorausgesetzt) dargelegt wird, ob das aus dieser Studie abzuleitende **tatsächliche** Risiko (der Fliegensitzpunkt) 0.4, 0.7 oder aber gar 2.4 betragen könnte. Auch für NNT sind Konfidenzintervalle zur Interpretation unerlässlich.

Noch ein letztes Wort zur Fliegenanalogie: Es geht also bei Aussagen aus klinischen Studien darum, auf der Basis der erzielten Ergebnisse eine Aussage über den theoretischen „wahren“ Effekt zu machen, sich ihm möglichst anzunähern und quantitativ zu beschreiben, also zu schätzen. Der „wahre“ Effekt selbst ist nicht beobachtbar! Die Möglichkeit eines Abgleichs zwischen dem „wahren“ Effekt und den in einer Studie erzielten Ergebnissen besteht nicht. Im Gegensatz zum Fliegenbeispiel

Tab. 47.1: Konfidenzintervalle (KI) für eine einzelne Rate

Sensitivität	Fallzahl*	Untere Grenze 95%-KI	Obere Grenze 95%-KI
80%	10	49%	94%
80%	50	67%	89%
80%	100	71%	87%
80%	1000	77%	82%

* **Achtung:** Für diese Rate ist dies, nicht zwangsläufig die Gesamtfallzahl der Studie, die Zahl der Studienteilnehmer mit der gesuchten Krankheit.

Tab. 47.2: Konfidenzintervalle (KI) für Effektmaße

Mortalität Behandlung	Mortalität Kontrolle	Relatives Risiko (RR)	Untere Grenze KI	Obere Grenze KI	NNT	Untere Grenze KI	Obere Grenze KI
2/20	4/20	0.5	0.1	2.4	10	NNH 3	8
6/60	12/60	0.5	0.2	1.25	10	NNH 5	38
50/500	100/500	0.5	0.37	0.69	10	7	18

Anmerkung:

Man kann unterschiedlicher Meinung darüber sein, ob bei einem nicht-signifikanten Studienergebnis überhaupt NNTs berechnet werden sollten. Einerseits führt dies zu merkwürdig erscheinenden Ergebnissen (was an der Berechnungsmethode liegt). Andererseits stellt sich die Frage, ob für dieses Maß, das den Behandlungseffekt quantifiziert, nicht zu nächst sichergestellt sein muss, dass der Behandlungseffekt überhaupt existiert. In der Tabelle wird dem Vorschlag von Altman gefolgt, in diesen Situationen als untere Grenze des KI die NNH (number needed to harm s. Glossar) anzugeben. Man beachte, dass zwischen den beiden angegebenen Grenzen des KI hier nicht die 0 liegt, sondern ∞! Die folgende Abbildung illustriert die Situation.

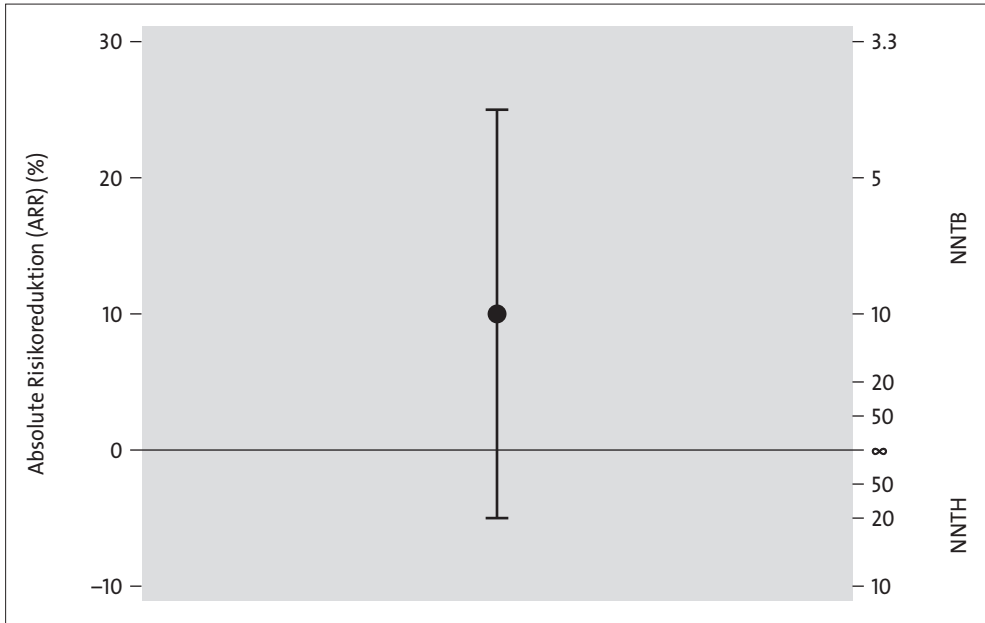


Abb 47.2: Verhältnis zwischen der Absoluten Risikoreduktion (ARR) und der „Number needed to treat“ aus: [2] (NNTB – Number needed to treat (benefit = Nutzen); NNTH – Number needed to treat (harm = Schaden))

gibt es in der klinischen Forschung nicht einmal eine Fliege, ein Ziel also, welches man zuverlässig anpeilen kann und welches sich **erst dann** der Beobachtung entzieht.

Das Prinzip des Schätzens lässt sich für jedes quantitative Maß, also auch für den Unterschied zwischen zwei Behandlungen (ausgedrückt in Relativen oder Absoluten Risiken) verwenden. Der Gebrauch von Konfidenzintervallen erlaubt dabei eine weitere Aussage. Es geht in solchen Studien (Stichwort: Testen) ja immer auch darum, die Aussage zu machen, ob überhaupt ein Unterschied zwischen zwei Therapien besteht bzw. ob eine Therapie einer anderen oder einem Placebo überlegen ist. Die Gleichheit von zwei Therapien lässt sich in den üblichen Effektmaßen einfach ausdrücken, bei Differenzen mit dem Wert 0, bei Quotienten mit dem Wert 1. Ist bei einer Studie im Konfidenzintervall für ein Absolutes Risiko die 0, im Konfidenzintervall für ein Relatives Risiko die 1 enthalten, so ist dies gleichbedeutend mit der Aussage, dass nicht eindeutig zwischen Gleichheit der Be-

handlungen und Überlegenheit einer Behandlung unterschieden werden kann (beide „Fliegensitzorte“ sind mit dem konkreten Studienergebnis vereinbar).

Dieses Ergebnis entspricht einem „nicht signifikanten“ Studienergebnis: Es ist nicht eindeutig zu sagen, dass sich das Studienergebnis von der Aussage „zwischen den beiden Behandlungen besteht kein Unterschied“, der sog. Nullhypothese, unterscheidet. Umgekehrt wäre ein Ergebnis dann als signifikant zu bezeichnen, wenn das Konfidenzintervall für eine Differenz die 0 **nicht** enthält, beide Grenzen also negative oder positive Vorzeichen haben bzw. bei einem relativen Effektmaß beide Grenzen kleiner oder beide Grenzen größer als 1 sind. In dieser Weise sind Aussagen aus Konfidenzintervallen mit Aussagen aus (den entsprechenden) statistischen Tests völlig äquivalent.

Im Gegensatz zum alleinigen Testen fällt jedoch beim Schätzen mit Konfidenzintervallen zusätzlich die quantitative Information an, die einerseits eine Bewertung im Schwarz-

Weiß-Bild signifikant/nicht signifikant erlaubt, andererseits aber zusätzlich zur Frage der Signifikanz die Bewertung der Relevanz ermöglicht. Nehmen wir an, man würde in einer klinischen Anwendungssituation aufgrund von Kosten oder bestehender Risiken nur Behandlungen mit einer NNT von höchstens 50 für vertretbar halten. In diesem Fall ist es für die Bewertung einer neuen Behandlung zwar nicht uninteressant, ob sie sich signifikant von einer anderen Behandlung unterscheidet. Auch das in einer Studie errechnete, also geschätzte NNT von z.B. 34 ist für eine Bewertung der Behandlung nicht ausreichend. Erst das Konfidenzintervall für das NNT von 34, das in diesem Beispiel von 13 bis 87 reichen soll, erlaubt die Aussage, dass die infrage stehende Behandlung zwar besser ist als eine Vergleichstherapie, aber die Relevanzgrenze von 50 nicht sicher (obere Grenze des Konfidenzintervalls 87) unterschreitet, seine uneingeschränkte Anwendung also möglicherweise nicht gerechtfertigt ist (zur Erinnerung: alle diese Aussagen gelten so nur, wenn die Abwesenheit von Bias vorausgesetzt werden kann).

Ein abschließendes Wort zur Studiengröße (n in der Konfidenzintervall-Formel): Ziel einer Studie ist es, eine für die Fragestellung ausreichend präzise Antwort zu geben. Wesentliche (in praxi oft einzige) Steuerungsmöglichkeit für diese Präzision ist die Fallzahl. Die Fallzahl muss also so hoch sein, dass z.B. die Sensitivität eines diagnostischen Tests mit für die praktische Entscheidungsfindung ausreichender Genauigkeit angegeben werden kann. Ein Wert von 75% mit einem Konfidenzintervall von 35% bis 97% (für $n=8$) ist belanglos. Genauso muss die Fallzahl einer vergleichenden Studie ausreichend hoch sein, damit gute Chancen bestehen, zwischen den Hypothesen „Es besteht kein Unterschied zwischen den beiden Therapien.“ und „Eine Therapie ist besser/ schlechter als die andere.“ zu unterscheiden (dies bezeichnet man als die „Power“ einer Studie).

Aus Aufwands- und auch aus ethischen Gründen ist es erforderlich, keine Patienten überflüssigerweise in eine Studie einzubeziehen. Daher muss die Fallzahl einer Studie genau so gewählt werden, dass gerade die erforderliche Präzision erreicht wird. Dieses Ziel wird – ohne hier auf weitere Details einzugehen – durch eine sog. Fallzahlplanung einer Studie erreicht, die für jede Studie erforderlich ist. Wie der Begriff deutlich macht, handelt es sich dabei um ein Instrument für die **Planung** einer Studie. Für die **Bewertung** von Studien hat die Fallzahlplanung nur eine untergeordnete Bedeutung. Hier stehen die Ergebnisse im Vordergrund und die Bewertung kann und muss mit den beiden Bewertungskriterien „erfolgreich vermiedenes Bias“ und „ausreichend präzise Aussage“ erfolgen.

Literatur

- [1] Schrader J et al., Morbidity and Mortality After Stroke, Eprosartan Compared with Nitrendipine for Secondary Prevention: principal results of a prospective randomized controlled study (MOSES). *Stroke*. (2005), 36, 1218–1226
- [2] Altman DG, Confidence intervals for the number needed to treat. *BMJ*. (1998), 317, 1309–1312
- [3] Fletcher RH et al. (2006) *Klinische Epidemiologie: Grundlagen und Anwendung*. Deutschsprachige Ausgabe. Huber, Bern
- [4] Bland M, Peacock J (2002) *Statistical questions in evidence-based medicine*. Oxford Univ. Pr., Oxford
- [5] Trampisch HJ, Windeler J (2000) *Medizinische Statistik*. Springer, Berlin
- [6] Feinstein AR (1985) *Clinical epidemiology*. Saunders, Philadelphia
- [7] Bland M (1990) *An introduction to medical statistics*. Oxford Med. Publ., Oxford
- [8] Altman DG et al. (2002) *Statistics with confidence*. BMJ Books, London
- [9] Kramer MS (1988) *Clinical epidemiology and biostatistics*. Springer, Berlin
- [10] Haynes RB et al. (2005) *Clinical epidemiology. How to do clinical practice research*. Lippincott Williams and Wilkins, London

48 Checklisten zur Bewertung von klinischen Studien und Leitlinien

Guido Lerzynski, Hilke Bertelsmann, Regina Kunz

In den letzten Jahren hat der Anteil von systematischen Übersichtsarbeiten an den veröffentlichten wissenschaftlichen Publikationen zugenommen. Ein Grund für diese Entwicklung liegt in der damit möglichen Zusammenstellung gesammelter wissenschaftlicher Erkenntnisse zu einem bestimmten Thema. Für die Erstellung einer systematischen Übersichtsarbeit ist es erforderlich, dass Einzelstudien bestimmte methodische Kriterien erfüllen. Die Festlegung auf diese Kriterien, die in einer Einzelstudie von Autoren erwartet werden, hat zu der Entwicklung von Checklisten geführt. Diese waren zunächst als Hilfestellung für Autoren zur standardisierten Studiendokumentation gedacht. Aus dieser Motivation heraus sind beispielsweise die Bewertungsinstrumente der CONSORT- und STARD-Gruppe entstanden, die die Berichtsqualität von therapeutischen bzw. diagnostischen Studien standardisieren.

Daneben haben sich weitere Bewertungsinstrumente entwickelt, mit denen die Glaubwürdigkeit und die Ergebnissicherheit von Studien im Rahmen der klinischen Epidemiologie analysiert und bewertet werden konnten. Zu diesen Instrumenten zählen z.B. das Evidence based Medicine Toolkit oder das AGREE-Instrument. Mit der fortschreitenden Anwendung der klinischen Epidemiologie in der Praxis hat sich der Einsatz von Checklisten zur Bewertung der Studienqualität als ein hilfreiches Mittel zur Erfassung der wesentlichen methodischen wie inhaltlichen Kriterien einer Studie durchgesetzt, sodass auch die ursprünglich für die

Verfasser von Studien gedachten Instrumente im Rahmen der klinischen Epidemiologie für die Analyse von Studien genutzt werden.

Mithilfe gezielter Fragen, die auf den Kriterien der klinischen Epidemiologie basieren, kann eine Studie auf die Validität der Ergebnisse überprüft werden. So kann eine kritische Beleuchtung zentraler Aspekte wie die Festlegung auf die zu untersuchende Studienpopulation oder die Erfassung der zu messenden Endpunkte eine Gesamtbewertung erleichtern und gegebenenfalls Limitationen in der Übertragbarkeit der Ergebnisse aufzeigen.

Darüber hinaus kann eine systematisierte Studienauswertung anhand von Fragebögen die Analyse einzelner Studien vereinfachen und vorhandene methodische Probleme zur Diskussion stellen. Grundsätzlich sieht das Bewertungsverfahren anhand von Checklisten in einem ersten Schritt die Abbildung des Studieninhalts vor, an die sich die kritische Bewertung anschließt.

Mittlerweile gibt es eine Fülle von Instrumenten, von denen im Folgenden einige exemplarisch skizziert werden sollen. Dabei werden zunächst vier Instrumente vorgestellt, die als Hilfestellung für Autoren zur Verbesserung der Berichtsqualität von Studien und Leitlinien entwickelt wurden. Daran schließen sich fünf Bewertungsinstrumente an, die im Rahmen der klinischen Epidemiologie den Nutzern von Literatur dienen sollen, die Validität von Studien und Leitlinien zu erheben.

48.1 Bewertungsinstrumente, die primär für Autoren zur Verbesserung der Berichtsqualität von Studien entwickelt wurden

CONSORT (Consolidated standards of reporting trials) [1–3]

Das CONSORT-Instrument war ursprünglich als eine Hilfe zur Darstellung von Studienergebnissen in einem randomisiert-kontrollierten Design gedacht. Auf den Grundprinzipien der klinischen Epidemiologie basierend, beinhaltet der Fragebogen 22 Einzelfragen und ein Fluss-Schema, das den Verlauf der Studienteilnehmer innerhalb der Studie – vom ersten Screening bis zur Endpunkterhebung – darstellen soll. Dabei sollen Informationen aus der Rekrutierungsphase der Studie, der Interventionszuordnung, der Nachbeobachtung und der Auswertung dargestellt werden, die unter anderem eine bessere Beurteilung des Intention-to-treat-Ansatzes der Autoren ermöglichen. Die CONSORT-Checkliste ist in mehreren Sprachen verfügbar und wurde von den meisten bedeutenden medizinischen Zeitschriften als Referenzstandard für die Berichtsqualität von randomisierten kontrollierten Studien (RCTs) akzeptiert. Neben der Funktion eines Leitfadens für Studienautoren wird die CONSORT-Checkliste zunehmend auch zur Bewertung der Studienqualität genutzt.

STARD (Standards for Reporting of Diagnostic Accuracy) [4]

Neben den therapeutischen Studien hat in den letzten Jahren die Publikationsdichte von diagnostischen Studien zugenommen. Dabei war die Berichtsqualität oft durch eine recht heterogene Abbildung der gesammelten Daten gekennzeichnet. Die Festlegung und Standardisierung der Berichtsqualität diagnostischer Studien ist das Ziel der STARD-Checkliste. Neben 25 Einzelfragen zu der untersuchten Population, der diagnostischen

Intervention und zu den Ergebnissen gehört zu der Checkliste ein Flussdiagramm, das die Anzahl der Teilnehmer je diagnostischem Verfahren gegenüberstellen soll. Im Mittelpunkt der Analysen steht die Aussagevalidität zur diagnostischen Genauigkeit von Prüfmethoden im Vergleich zu Referenzmethoden. Neben der Anwendung für Autoren von diagnostischen Studien findet der STARD-Bogen zunehmend Verwendung bei der Bewertung der Studienqualität diagnostischer Studien.

QUADAS (Quality assessment of studies of diagnostic accuracy included in systematic reviews) [5]

Ähnlich wie das STARD-Instrument gibt der QUADAS-Bogen eine Hilfestellung zur Darstellung diagnostischer Studien. Im Gegensatz zum STARD-Bogen steht hier die Zusammenfassung einzelner diagnostischer Studien in einem systematischen Review im Vordergrund. In den insgesamt 14 Einzelpunkten des Instruments, die Informationen zum eingeschlossenen Patientenkreis, zum Prüf- und Referenztest, zur Testdurchführung und zur Gruppe der ausgeschiedenen Teilnehmer erheben, liegt ein weiterer Schwerpunkt in der Erfassung der für diagnostische Studien und systematische Reviews besonders relevanten Formen von Bias.

QUORUM (Quality of Reporting of Meta-analyses) [6]

In dem Sinne, wie der QUADAS-Bogen eine Hilfestellung für Autoren in der Zusammenfassung einzelner diagnostischer Studien zu einem systematischen Review sein soll, dient das QUORUM-Instrument Verfassern von quantitativen Übersichtsarbeiten therapeutischer Studien (Meta-Analyse). 21 Einzelbereiche hinterfragen die systematische Literaturrecherche, die Auswahl der Selektionskriterien, die Glaubwürdigkeit und Merkmale der Einzelstudien, die Art der Datenextrakti-

on und die Art der quantitativen Zusammenfassung der Einzelstudien. Das abschließende Flussdiagramm dient der transparenten Darstellung der eingeschlossenen und ausgeschlossenen RCTs einschließlich der entsprechenden Begründung für den Ein- bzw. Ausschluss.

48.2 Bewertungsinstrumente, die primär für die Nutzer von evidenzbasierten Studienergebnissen entwickelt wurden

AGREE (Appraisal of Guidelines Research & Evaluation) [7]

Das AGREE-Instrument ist als Beurteilungsinstrument für die Nutzer von evidenzbasierten Ergebnissen aus Leitlinien entwickelt worden. In 23 Einzelaspekten werden sechs Dimensionen einer Leitlinie erfasst. Dazu zählen der Geltungsbereich und Zweck einer Leitlinie, die Beteiligung von Interessengruppen bei der Erstellung des Dokuments, die methodische Exaktheit der Leitlinie, die Klarheit und die Gestaltung, die Anwendbarkeit und die redaktionelle Unabhängigkeit der Leitlinie. Um die Zuverlässigkeit der Bewertung einer Leitlinie zu erhöhen, wird die Bewertung einer Leitlinie mittels des AGREE-Instruments durch mindestens 2 Gutachter empfohlen.

DELBI (Deutsches Leitlinien Bewertungsinstrument) [8]

Das für den deutschen Sprachraum zwischen 2003 und 2005 entwickelte Instrument soll die Bewertung der methodischen Qualität medizinischer Leitlinien ermöglichen. Es ist eine Weiterentwicklung des AGREE-Instruments unter Berücksichtigung der Erfahrungen des Internationalen Leitlinien-Netzwerks (G-I-N), der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) und des Ärztlichen Zentrums für Qualität in der Medizin (ÄZQ).

CASP (Critical appraisal skills programme) [9]

Das „Critical Appraisal Skills Programme (CASP)“ ist ein Programm des Public Health-Bereichs des britischen Gesundheitssystems (NHS), das für Beschäftigte im Gesundheitswesen entwickelt wurde. Seit 1993 wurden dort evidenzbasierte Programme und Anwendungen entwickelt, die allen Arten von Gesundheitsberufen helfen sollen, ihre Fertigkeiten im Auffinden und Bewerten von relevanten Studien sowie im Anwenden auf klinisch relevante Fragestellungen zu verbessern.

Das Programm hat unter diesem Aspekt Bewertungsbögen für RCTs, systematische Reviews, qualitative und diagnostische Studien, Kohorten- und Fall-Kontroll-Studien sowie ökonomische Analysen entwickelt, die die Bewertung der Evidenz erleichtern sollen.

Evidence based Medicine Toolkit [10]

Das Evidence based tool kit umfasst ähnlich wie das CASP-Programm eine Sammlung von Bewertungsinstrumenten, um für eine umschriebene Fragestellung die relevante Evidenz zu finden, zu bewerten und umzusetzen. Die einzelnen Checklisten entstammen der Users' Guides Serie, die im Journal of the American Medical Association (JAMA) von der Evidence Based Medicine Working Group entwickelt wurde.

ASSERT (A Standard for the Scientific and Ethical Review of Trials) [11]

Das ASSERT-Instrument stellt die Bewertung ethischer Fragestellungen von RCTs in den Vordergrund. Die Checkliste untersucht, inwieweit in den Studien eine ethische Herangehensweise berücksichtigt wurde. Untersuchte Aspekte sind der wissenschaftliche und soziale Wert einer Erhebung, die wissenschaftliche Gültigkeit, die faire Teilnehmerrekrutierung sowie die Würde und der Respekt gegenüber den in die Studie eingeschlossenen Teilnehmern.

Tab. 48.1: Übersicht ausgewählter Bewertungsinstrumente (Stand: Juni 2006)

Bewertungsinstrument	Studientyp	Literaturangaben, Online-Verfügbarkeit
CONSORT (Consolidated Standards of Reporting Trials)	Randomisierte kontrollierte Studien (RCTs)	http://www.consort-statement.org/
STARD (Standards for Reporting of Diagnostic Accuracy)	Diagnostische Studien	http://www.consort-statement.org/stardstatement.htm
QUADAS (Quality assessment of studies of diagnostic accuracy included in systematic reviews)	Diagnostische Studien	BMC Med Res Methodol. 2003; 3, 25
QUORUM (Quality of Reporting of Meta-analyses)	Systematische Übersichten (SRs), Meta-Analysen	The Lancet 1999 Nov 27; 354(9193):1896–900
DELBI (Deutsches Leitlinien-Bewertungsinstrument)	Leitlinien	www.delbi.de
AGREE (Appraisal of Guidelines Research & Evaluation)	Leitlinien	http://www.g-i-n.net
CASP (critical appraisal skills programme)	RCTs, SRs, Fall-Kontroll-Studien, Kohortenstudien, diagnostische Studien, qualitative Studien, ökonomische Evaluationsstudien	http://www.phru.nhs.uk/casp/critical_appraisal_tools.htm
Evidence based Medicine Toolkit (Evidence Based Medicine Working Group, Canada)	SRs, Leitlinien, therapeutische, diagnostische und prognostische Studien sowie ökonomische Analysen	http://www.med.ualberta.ca/ebm/ebm.htm
ASSERT (A Standard for the Scientific and Ethical Review of Trials)	Bewertung ethischer Aspekte im Studiendesign	http://www.assert-statement.org/

Literatur

- [1] Freemantle N et al., CONSORT: an important step toward evidence-based health care. Consolidated Standards of Reporting Trials. *Ann Intern Med.* (1997), 126, 81–83
- [2] Begg C et al., Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA.* (1996), 276, 637–639
- [3] Altman DG, Better reporting of randomised controlled trials: the CONSORT statement. *BMJ.* (1996), 313, 570–571
- [4] Bossuyt PM et al., The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem.* (2003), 49, 7–18
- [5] Whiting P et al., The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* (2003), 3, 25
- [6] Clarke M, The QUORUM statement. *Lancet.* (2000), 355, 756–757
- [7] The AGREE Collaboration, Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Saf Health Care.* (2003), 12, 18–23. <http://www.g-i-n.net>
- [8] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI). Fassung 2005/2006. *Z Arztl Fortbild Qualitatssich.* (2005), 99, 468–519. <http://www.delbi.de>
- [9] NHS, Critical Appraisal Skills Programme (CASP) and Evidence-based Practice. <http://www.phru.nhs.uk/casp/casp.htm> (14.12.06)
- [10] EBM toolkit. <http://www.med.ualberta.ca/ebm/ebm.htm> (14.12.06)
- [11] The ASSERT statement, A standard for the scientific and ethical review of trials. <http://www.assert-statement.org> (14.12.06)

49 Glossar zur Evidenzbasierten Medizin

Regina Kunz, Dagmar Lühhmann, Jürgen Windeler, Monika Lelgemann, Norbert Donner-Banzhoff

Absolute Risiko-Reduktion (ARR)

Effektmaß für dichotome Endpunkte; beschreibt die absolute Differenz der Rate an ungünstigen Ereignissen in der experimentellen Gruppe (E) im Vergleich zur Kontrollgruppe (K), wenn die experimentelle Behandlung wirksam ist ($ARR = K - E$). Der Kehrwert der ARR ergibt die *Number needed to treat* ($1/ARR = NNT$).

Absolute Risiko-Zunahme (ARI, absolute risk increase)

Beschreibt die absolute Differenz der Rate an ungünstigen Ereignissen in der experimentellen Gruppe (E) im Vergleich zur Kontrollgruppe (K), wenn die experimentelle Behandlung schlechter ist ($ARI = |K - E|$). Der Kehrwert der ARI ergibt die *Number needed to harm* ($1/ARI = NNH$).

Ausschlusskriterien

Bei klinischen und epidemiologischen Studien wird die Studienpopulation über explizite Ausschluss- und *Einschlusskriterien* definiert. Ausschlusskriterien beschreiben negativ formulierte Eigenschaften, die potenzielle Kandidaten von einer Studienteilnahme ausschließen. Sie leiten sich zum großen Teil aus der klinischen oder epidemiologischen Studienfragestellung ab (z.B. Ausschluss von besonderen Risikopatienten), zum Teil auch aus ethischen und rechtlichen Vorgaben. Letztere kommen vor allem bei Interventionsstudien zum Tragen.

Behandlungsgruppe (Synonyme:

Experimentalgruppe, Interventionsgruppe, Verumgruppe)

Der Anteil an der Studienpopulation in einer kontrollierten Interventionsstudie, der die zu untersuchende Behandlung/Intervention erhält; die Resultate der Behandlungsgruppe werden mit denen der *Kontrollgruppe* verglichen, um Aussagen über Größe und Richtung von Interventionserfolgen zu erhalten. (s. auch *Kontrollgruppe*; *Interventionsgruppe*; *Experimental event rate*; *Randomisierung*; *randomisierte kontrollierte Studie*)

Beobachtervariabilität (observer variability)

Schwankungen von Mess- oder Untersuchungsergebnissen, die mit der Person des Beobachters in Zusammenhang stehen. Sie ist umso größer, je geringer die Reliabilität des verwendeten Messinstruments ist (z.B. Auskultation von Herztönen); Befunderhebung und/oder -interpretation sind ggfs. beeinflussbar durch die subjektive Wahrnehmung, Erfahrung und Ausbildung, evtl. auch durch Wunschdenken des Untersuchers. Es werden unterschieden: 1. Die **Intraobserver-Variabilität** bezeichnet die Schwankungen der Ergebnisse, wenn eine Person den gleichen (unveränderten) Untersuchungsgegenstand mehrfach untersucht. 2. Die **Interobserver-Variabilität** bezeichnet die Schwankungen der Ergebnisse, wenn mehrere Personen denselben Untersuchungsgegenstand untersuchen.

Beobachtungsgleichheit (consistency of observation)

Anforderung an die Durchführung von kontrollierten Studien, alle Studienteilnehmer gleichartig zu beobachten, um Ergebnisverzerrungen durch unterschiedliche Beobachtungsmodalitäten zu verhindern (s. auch *Bias*).

Bias (systematischer Fehler)

Tendenz der Studienergebnisse, systematisch von den „wahren“ Ergebnissen abzuweichen. Bias führt entweder zu einer Über- oder Unterschätzung der wahren Wirkung einer Maßnahme oder Exposition. Die Ursachen dafür liegen vor allem im Design und der Durchführung der Studie und führen zu systematischen Unterschieden zwischen den Vergleichsgruppen, z.B. bei der Auswahl der Teilnehmer (Selektions-Bias), der Erhebung der Endpunkte (observation bias oder Beobachtungs-Bias) oder dem Verlust von Teilnehmern in der Studie (attrition bias oder Verschleiß-Bias). Ergebnisse aus Studien mit geringem Risiko für Bias werden als valide angesehen.

Carry-over-Effekt

Störender Effekt in *Interventionsstudien* mit *Cross-over-Design*, wenn die Behandlung der 1. Phase über das Ende dieser Phase hinaus und in die zweite Phase hinein wirkt. So kann das Ergebnis der Behandlung in der 2. Phase beeinflusst und der Vergleich verzerrt werden. Zur Vermeidung dieses Effekts kann vor die 2. Studienphase eine so genannte Wash-out-Phase gesetzt werden, in der die Wirkungen vorangegangener Therapien abklingen.

Confounding

Confounding liegt vor, wenn ein Faktor (Confounder), der nicht direkt Gegenstand der Untersuchung ist, sowohl mit der Intervention/Exposition als auch mit der Zielgröße assoziiert ist und dadurch bei Aussagen

über die Beziehung zwischen Intervention/Exposition und Zielgröße „Verwirrung“ stiftet. Häufige Confounder sind z.B. Alter, Geschlecht oder Rauchen. Confounding lässt sich durch ein entsprechendes Studiendesign (z.B. Randomisierung oder Matching) oder durch die Anwendung bestimmter statistischer Verfahren bei der Analyse (Stratifizierung, multivariate Analyse) kontrollieren.

Control event rate (CER; Ereignisrate in der Kontrollgruppe)

Anteil der Teilnehmer in der Kontrollgruppe, die in einem definierten Zeitraum ein Ereignis oder einen Endpunkt erleiden. Die Ereignisrate in der Kontrollgruppe wird zur Berechnung der *Absoluten Risiko-Reduktion* und *Relativen Risiko-Reduktion* benötigt.

Critical Appraisal

In der *Evidenzbasierten Medizin*: kritische Bewertung von wissenschaftlichen Studien mit dem Ziel, zu entscheiden, ob diese geeignet sind, zur Lösung eines spezifizierten Problems aus der medizinischen Versorgung (z.B. beste Behandlungsform für den eigenen Patienten) beizutragen; geprüft werden: 1. *Validität*; 2. klinische Relevanz; 3. Übertragbarkeit auf den gegebenen Entscheidungskontext (s. auch *Bias*).

Critically Appraised Topic (CAT)

In der *Evidenzbasierten Medizin* (EbM) ein nach EbM-Methodik gelöstes Problem, das die präzise formulierte Fragestellung, eine kurze Beschreibung der Informationsrecherche, die strukturierte Kurzfassung einer geeigneten, kritisch bewerteten Studie und die aus den Ergebnissen abgeleitete klinische Schlussfolgerung enthält; unter didaktischen Gesichtspunkten eignet sich die Erstellung von CATs zum Erlernen und Vertiefen der EbM-Methoden. Im Internet stehen freie Programme, so genannte CAT-Maker, zur Verfügung. Wenn eine CAT-Sammlung als Informationsquelle weiterverwendet werden soll,

sind 3 Probleme zu bedenken: 1. CATs können fehlerhaft sein, ein Peer-Review ist daher unverzichtbar. 2. CATs geben meist nur eine einzige Quelle zum klinischen Problem wieder und sind fast nie repräsentativ für die gesamte, zum Thema publizierte Literatur. 3. Je nach Thema veralten CATs schnell.

Cross-over-Design

In Cross-over-Studien wird die Wirksamkeit zweier Behandlungsformen verglichen, indem diese zeitlich versetzt den gleichen Probanden verabreicht werden. Die typische Cross-over-Studie hat mindestens zwei Studienphasen. In der ersten Studienphase erhalten die Probanden des ersten Studienarmes zunächst Therapie A, die Probanden des zweiten Studienarmes Therapie B. Nach einer definierten Behandlungsdauer wird zur 2. Studienphase gewechselt: Studienarm eins erhält nun Therapie B und Studienarm zwei Therapie A. Die Wirksamkeit der Therapien kann am Ende der Studie sowohl im Intra- als auch im Intergruppenvergleich bestimmt werden. Effizientes Studiendesign, da Variabilität („Rauschen“) vermindert wird; nicht geeignet bei progredienten Erkrankungen.

Diskrete Variablen

Diskrete Variablen weisen im Gegensatz zu kontinuierlichen Variablen nur eine begrenzte Zahl eindeutig voneinander abgrenzbarer Zustände auf (z.B. Augenfarbe: blau, grau, braun, grün). Eine Sonderform sind dichotome Variablen, die lediglich zwei Alternativen aufweisen, z.B. Raucher/Nichtraucher, lebend oder tot, Test-positiv oder Test-negativ.

Dichotome Variable

s. Diskrete Variable

Entscheidungsbaum (decision tree)

Darstellung von Entscheidungsregeln, welche die logische Abhängigkeit von hierarchisch angeordneten Ereignissen in Form ei-

nes Baumdiagramms veranschaulicht. Die Wahlknoten an den Astverzweigungen sind mit einer bestimmten (angenommenen oder bekannten) Wahrscheinlichkeit für den Nutzen der Intervention versehen. Von diesen gehen weitere Äste mit neuen Ergebnissen und entsprechenden Wahrscheinlichkeiten ab. Die Wahrscheinlichkeit des Nutzens einer Maßnahme kann so abgeleitet werden, die Validität der Aussage ist aber natürlich abhängig von der Güte der eingesetzten Informationen. Entscheidungs bäume sind Entscheidungshilfen zur Visualisierung komplexer Abhängigkeitsstrukturen, zur Identifizierung aller Entscheidungsmöglichkeiten und deren Ergebnissen sowie für die Lösung von Wahrscheinlichkeitsaufgaben.

Effektmaß

Maßzahl, um die Stärke eines Effekts zu quantifizieren. Gebräuchliche Effektmaße für dichotome Endpunkte sind das *relative Risiko* (RR) oder die *Odds-Ratio* (OR), gebräuchliche Effektmaße für kontinuierliche Endpunkte sind in Einzelstudien die (standardisierte) mittlere Differenz und in Meta-Analysen die *gewichtete mittlere Differenz* (Weighted mean difference, WMD).

Effectiveness („Wirksamkeit unter Alltagsbedingungen“)

Beschreibt die Wirksamkeit einer Maßnahme unter Routinebedingungen. Im Gegensatz zur *Efficacy* („Wirksamkeit unter Idealbedingungen“) untersuchen Effectiveness-Studien die Frage: Wirkt die Maßnahme unter den Bedingungen der Routineversorgung?

Efficacy („Wirksamkeit unter Idealbedingungen“)

Beschreibt die Wirksamkeit einer Maßnahme unter Idealbedingungen. Efficacy-Studien sollen eine hohe interne Validität haben, die Ergebnisse sind jedoch möglicherweise nur bedingt auf die Routineversorgung übertragbar.

Einschlusskriterien

Bei klinischen und epidemiologischen Studien wird die Studienpopulation über explizite Einschluss- und *Ausschlusskriterien* definiert. Einschlusskriterien beschreiben positiv formulierte Eigenschaften, welche die potenziellen Studienteilnehmer erfüllen müssen. Sie leiten sich einerseits ab aus der klinischen oder epidemiologischen Studienfragestellung, andererseits aus rechtlichen und ethischen Vorgaben. Typische Einschlusskriterien sind das Vorliegen einer bestimmten Erkrankung, gegebenenfalls spezifiziert nach Schweregrad, bisherigem Verlauf und Vorbehandlungen oder die Zugehörigkeit zu einer definierten Altersgruppe.

Evidenz (evidence)

Der Begriff „Evidenz“ im Kontext der Evidenzbasierten Medizin leitet sich vom englischen Wort „evidence“ = Nach-, Beweis ab und bezieht sich auf die Informationen aus klinischen Studien, die einen Sachverhalt erhärten oder widerlegen.

Evidenzbasierte Medizin (Evidence-Based Medicine, EbM)

Unter Evidenzbasierter Medizin (EbM) oder evidenzbasierter Praxis im engeren Sinne versteht man eine Vorgehensweise des medizinischen Handelns, individuelle Patienten auf der Basis der besten zur Verfügung stehenden Daten zu versorgen. Diese Technik umfasst die systematische Suche nach der relevanten Evidenz in der medizinischen Literatur für ein konkretes klinisches Problem, die kritische Beurteilung der Validität der Evidenz nach klinisch-epidemiologischen Gesichtspunkten; die Bewertung der Größe des beobachteten Effekts sowie die Anwendung dieser Evidenz auf den konkreten Patienten mithilfe der klinischen Erfahrung und der Vorstellungen der Patienten. Ein verwandter Begriff ist die **evidenzbasierte Gesundheitsversorgung (Evidence-Based Health Care)**, bei der die Prinzipien der EbM

auf alle Gesundheitsberufe und alle Bereiche der Gesundheitsversorgung, einschließlich Entscheidungen zur Steuerung des Gesundheitssystems, angewandt werden.

Experimentalgruppe

s. Behandlungsgruppe

Experimental event rate (EER; Ereignisrate in der experimentellen Gruppe)

Anteil der Teilnehmer in der experimentellen Gruppe einer klinischen Studie, die in einem definierten Zeitraum ein Ereignis oder einen Endpunkt erleiden. Die Ereignisrate (*Risiko*) in der experimentellen Gruppe wird zur Berechnung der *absoluten Risiko-Reduktion* und *relativen Risiko-Reduktion* benötigt.

Externe Validität (Übertragbarkeit, Anwendbarkeit)

Beschreibt die Übertragbarkeit von Studienergebnissen auf die Patienten in der Routineversorgung, d.h. auf Patienten jenseits des Studiensettings (s. a. *Interne Validität*).

Fall-Kontroll-Studie

Retrospektive Beobachtungsstudie, bei der eine Gruppe von Personen mit einer Zielerkrankung („Fälle“) und eine Gruppe von Personen ohne die Erkrankung („Kontrollen“) auf das Vorhandensein von Expositionsfaktoren (Risiko- oder protektive Faktoren) verglichen werden.

Gewichtete mittlere Differenz (Weighted mean difference, WMD)

Effektmaß für kontinuierliche Endpunkte (s. *Effektmaß*), die auf derselben Skala gemessen werden (z.B. Größe) zur Beschreibung des Gesamteffekts, wenn Studien in Meta-Analysen gepoolt werden. Dabei erhalten die Einzelstudien ein unterschiedliches Gewicht, um wichtige Studieninformationen, wie z.B. Größe der Studie oder Präzision des Effekts, zu berücksichtigen.

Hawthorne-Effekt/Hawthorne-Experiment

Phänomen, dass in Studien allein durch die Tatsache, dass Probanden vermehrte Aufmerksamkeit erfahren, Effekte erzeugt werden können. Erkenntnis stammt aus einem kontrollierten sozialwissenschaftlichen Experiment, mit dem 1927 in den Hawthorne-Werken (USA) überprüft werden sollte, inwieweit eine Verbesserung der äußeren Arbeitsbedingungen (Ergonomie, Lichtverhältnisse, Pausenzeiten, Länge des Arbeitstages) die Produktivität von Arbeiterinnen steigern kann. Vor, während und nach der Versuchsreihe fanden intensive Besprechungen mit den Probandinnen statt. Die Leistungen der Arbeiterinnen mit geänderten Arbeitsbedingungen wurden mit den Leistungen einer Kontrollgruppe verglichen. Überraschenderweise stieg die Produktivität in beiden Gruppen an, d.h. auch in der Kontrollgruppe, die nur vermehrte Aufmerksamkeit, aber keine verbesserten Arbeitsbedingungen erfuhr.

Heterogenität/Homogenität

In systematischen Reviews oder Meta-Analysen bezeichnet Homogenität bzw. Heterogenität, inwieweit die in den eingeschlossenen Studien gefundenen Effekte ähnlich (homogen) bzw. verschieden (heterogen) sind. Mit statistischen Heterogenitätstests kann festgestellt werden, ob die Unterschiede zwischen den Studien größer sind als zufallsbedingt zu erwarten wäre. Als Ursachen für Heterogenität kommen Unterschiede in den Patientencharakteristika, Intervention oder Endpunkte zwischen den Studien infrage, was aus klinischer Sicht beurteilt werden muss. Die Durchführung einer *Meta-Analyse* aus heterogenen Studien ist problematisch.

Homogenität

s. Heterogenität

Hypothese

Aussage, Behauptung oder Theorie, deren Zutreffen oder Nicht-Zutreffen mithilfe einer

Studie oder eines Experimentes geprüft werden soll.

Hypothesengenerierung

Prozess des Ableitens von *Hypothesen*, die in einer Studie oder mithilfe eines Experimentes überprüft werden können. Überprüfbare Hypothesen werden aus theoretischen Ableitungen, zufälligen oder systematischen Beobachtungen oder Pilotstudien gewonnen.

Interne Validität

Interne Validität bezeichnet das Ausmaß, in dem die Ergebnisse einer Studie die „wahren“ Effekt einer Intervention/Exposition wiedergeben, d.h. inwieweit die Studie frei von systematischen Fehlern (*Bias*) ist. Die interne Validität beruht auf der Integrität des Studiendesigns und ist Voraussetzung für die Anwendbarkeit der Studienergebnisse in der Routineversorgung. (s. auch *externe Validität*)

Intention-to-treat-Analyse (ITT)

Analyse-Technik, bei der die Patienten nach ihrer ursprünglichen Gruppenzuteilung analysiert werden, unabhängig davon, ob sie die zugeordnete (intendierte) Therapieform vollständig, partiell oder gar nicht erhalten haben. ITT kann heute als das Standard-Auswertungsverfahren für klinische Studien bezeichnet werden. Sie sichert die Strukturgleichheit von Interventions- und Kontrollgruppe.

Intervention

In klinischen Studien Bezeichnung für die Maßnahme, deren Wirksamkeit untersucht werden soll. (s. auch *Behandlungsgruppe; Kontrollgruppe*)

Interventionsgruppe

s. *Behandlungsgruppe*

Inzidenz

Die Inzidenz beschreibt die in einem bestimmten Zeitraum neu aufgetretene Anzahl

von Krankheitsfällen in einer definierten Population. (s. auch *Prävalenz*)

Klinische Studie

Unschärf definierter Begriff für eine Studie, in der eine Intervention an einer Gruppe von Patienten untersucht wird. Oberbegriff für unterschiedliche Studientypen, z.B. nicht kontrollierte, kontrollierte und randomisierte klinische Studien, auch für diagnostische, prognostische und ätiologische Fragestellungen.

Kohorte

Gruppe von Personen mit definierten Charakteristika (z.B. Gesundheitszustand), die in klinischen und epidemiologischen Studien beobachtet werden (s. Kohortenstudie); Einteilung: 1. geschlossene Kohorte: Alle Probanden werden zum gleichen Zeitpunkt in die Beobachtung eingeschlossen und verbleiben dort bis zum Eintritt des Zielereignisses oder bis zum Ende der Studie. 2. offene Kohorte: Personen werden nach und nach in die Beobachtung aufgenommen und/oder scheiden nach einer definierten Beobachtungsdauer wieder aus der Studie aus. 3. Sonderform Geburtskohorte: enthält alle Personen einer umschriebenen Region, die im gleichen, meist eng definierten Zeitraum geboren wurden.

Kohortenstudie

Vergleichende Beobachtungsstudie, in der Personen (Kohorte) mit bzw. ohne eine Intervention/Exposition (zu der sie nicht unbedingt von dem Studienarzt zugeteilt wurden) über einen definierten Zeitraum beobachtet werden, um Unterschiede im Auftreten der Zielerkrankung festzustellen. Kohortenstudien können prospektiv oder retrospektiv durchgeführt werden.

Konfidenzintervall (Vertrauensbereich, confidence interval – CI)

Bereich, in dem der „wahre“ Wert einer Messung (Effektgröße) mit einer bestimmten

Wahrscheinlichkeit erwartet werden kann (üblicherweise 95% → 95%-Konfidenzintervall). Die Effektgröße kann dabei z.B. ein Therapieeffekt, ein Risiko oder die Sensitivität eines diagnostischen Tests sein. Das Konfidenzintervall beschreibt die Unsicherheit über die Zuverlässigkeit der Aussage zur Effektgröße. Die Breite des Konfidenzintervalls hängt u.a. von der Zahl der in die Studie eingeschlossenen Patienten ab und wird mit zunehmender Patientenzahl enger, d.h. die Effektgröße kann präziser geschätzt werden.

Kontinuierliche/stetige Variablen

Im Gegensatz zu *diskreten Variablen* können kontinuierliche Variablen theoretisch eine unendlich große Zahl von Werten entlang eines Kontinuums annehmen. Körpergröße, Gewicht und viele Laborwerte sind kontinuierliche Variablen.

Kontrollgruppe

Anteil an der Studienpopulation in einer kontrollierten Interventionsstudie, der die zu untersuchende Behandlung/*Intervention* nicht erhält, aber ggfs. mit einer Placebointervention oder einer Standardbehandlung versorgt wird; die Resultate der Kontrollgruppe werden mit denen der *Behandlungsgruppe* verglichen, um Aussagen über Größe und Richtung von Interventionserfolgen zu erhalten. (s. auch *Behandlungsgruppe*; *Control Event Rate*; *Randomisierung*; *Hawthorne-Experiment*)

Likelihood-Ratio

s. *Wahrscheinlichkeitsverhältnis*

Matched-Pairs-Technik

Verfahren zur Herstellung von möglichst strukturgleichen Untersuchungsgruppen in kontrollierten klinischen und epidemiologischen Studien. „Strukturgleich“ bezieht sich auf alle Faktoren, die einen wesentlichen Einfluss auf die Zielgröße der Studie haben können. Für jedes Individuum der Untersuchungsgruppe wird ein in definierten Ein-

flussfaktoren (Matchingkriterien) entsprechendes Individuum in die Kontrollgruppe aufgenommen. Häufig verwendete Matchingkriterien sind Alter, Geschlecht, Ethnizität, Schweregrad der Erkrankung, Komorbidität und Risikofaktoren.

Matching („Anpassung“)

Bestreben, zwei Untersuchungsgruppen möglichst vergleichbar zu machen.

Beim **Matching nach Häufigkeiten** werden die Merkmalsträger in den zu vergleichenden Gruppen ausgezählt (Frequency Matching); z.B. sollen in einer Studie zur koronaren Herzkrankheit in beiden Gruppen gleich hohe Anteile von Frauen, Rauchern, Diabetikern oder Hypertonikern vertreten sein, um eine systematische Verzerrung der Ergebnisse zu verhindern.

Beim **individuellen Matching** sollen für jedes Individuum der Untersuchungsgruppe ein oder mehrere den Matchingkriterien entsprechende Individuen in die Kontrollgruppe aufgenommen werden. (s. auch *Matched-Pairs-Technik*)

Meta-Analyse

Statistisches Verfahren, um die Ergebnisse mehrerer Studien, welche die gleiche Frage bearbeiten, quantitativ zu einem Gesamtergebnis zusammenzufassen und dadurch die Aussagekraft (Genauigkeit der Effektschätzer) gegenüber Einzelstudien zu erhöhen. Meta-Analysen werden mit zunehmender Häufigkeit in *systematischen Reviews* eingesetzt. Allerdings beruht nicht jede Meta-Analyse auf einem *systematischen Review*.

Nachtest-Wahrscheinlichkeit einer Erkrankung (post-test probability)

Beschreibt die Wahrscheinlichkeit des Vorliegens einer Erkrankung unter Berücksichtigung der Ergebnisse eines diagnostischen Tests. Ist sie niedrig, wird die Spur meist nicht weiter verfolgt; ist sie hoch, wird in der Regel behandelt; in einem mittleren Wahr-

scheinlichkeits-Bereich ist weitere Diagnostik sinnvoll.

Negativer prädiktiver Wert

Anteil der Personen mit negativem Testergebnis, bei denen die gesuchte Krankheit tatsächlich nicht vorliegt. Dieser Wert hängt von der *Prävalenz* der Erkrankung in der untersuchten Gruppe ab.

Nullhypothese

Bei der Durchführung statistischer Signifikanztests wird die Hypothese aufgestellt, dass zwischen den verschiedenen Gruppen einer Studie kein Unterschied besteht. Aus statistischer Sicht ist z.B. eine Behandlung wirksam, wenn man durch den statistischen Test die Nullhypothese, dass es zwischen den Ergebnissen der experimentellen und der Kontrollgruppe keinen Unterschied gibt, verwerfen kann. (s. auch *statistische Signifikanz*)

Number needed to treat (NNT)

Klinisch intuitives Effektmaß für dichotome Endpunkte, um die Auswirkung einer Behandlung zu beschreiben. Gibt die Anzahl von Patienten wieder, die behandelt werden müssen, um 1 zusätzliches ungünstiges Ereignis zu verhindern. Die NNT wird als $1/ARR$ (s. a. *Absolute Risiko-Reduktion*) berechnet.

Number needed to harm (NNH)

Klinisch intuitives Effektmaß für dichotome Endpunkte, um die ungünstigen Auswirkung einer Behandlung zu beschreiben. Gibt die Anzahl an Patienten wieder, bei deren Behandlung mit einem zusätzlichen Fall unerwünschter Ereignisse/einer Komplikation gerechnet werden muss. Die NNT wird als $1/ARI$ (s. a. *absolute Risiko-Zunahme*) berechnet.

Odds (Chance)

Beschreibt in einer Gruppe das Verhältnis zwischen der Anzahl von Teilnehmern mit einem Endpunkt und der Anzahl von Teilnehmern ohne diesen Endpunkt. Wenn also

30 von 100 Teilnehmern den Endpunkt entwickeln (und 70 nicht), beträgt die Odds 30:70 oder 0.42 (beschreibt den gleichen Sachverhalt wie das *Risiko*).

Odds-Ratio (OR, Chancenverhältnis)

Effektmaß für *dichotome* Daten. Bezeichnet das Verhältnis (Ratio) der Odds, dass ein Ereignis oder Endpunkt in der experimentellen Gruppe eintritt, zu der Odds, dass das Ereignis in der Kontrollgruppe eintritt. Eine OR von 1 bedeutet, dass zwischen den Vergleichsgruppen kein Unterschied besteht. Bei ungünstigen Endpunkten zeigt eine OR < 1, dass die experimentelle Intervention wirksam ist, um die Odds für das Auftreten dieser ungünstigen Endpunkte zu senken. (s. auch *Relatives Risiko*)

Outcomes (Ergebnis, Endzustand)

Übergeordneter Begriff für die Auswirkungen von therapeutischen oder präventiven Maßnahmen und anderen Expositionen.

p-Wert

p-Werte (p von probability) beschreiben die Wahrscheinlichkeit für den beobachteten (oder einen noch extremeren) Effekt einer Studie und zwar unter der Bedingung, dass die *Nullhypothese* richtig ist. Je kleiner der Wert, desto deutlicher spricht das beobachtete Ergebnis gegen die Nullhypothese. Es ist eine Konvention, dass ein p-Wert gleich oder kleiner 0.05 als statistisch signifikant angesehen wird. Wenn die Signifikanz von Effekten interpretiert wird, sollten p-Werte immer im Zusammenhang mit *Konfidenzintervallen* verwendet werden.

Parallelgruppen-Studie

Klassisches Design von kontrollierten Interventionsstudien, bei denen die Studienarme zeitlich parallel laufen. (s. auch *Cross-over-Studie*)

Per-protocol-Analyse

Auswertungsmethode für (randomisierte kontrollierte) Interventionsstudien. Die Ergebnisse der Studienteilnehmer werden nur dann in der Auswertung berücksichtigt, wenn die Behandlungen (Verum- oder Kontrollbehandlung) protokollgemäß durchgeführt wurden. Die Per-protocol-Analyse erlaubt die Abschätzung des Wirkungspotenzials einer Behandlungsform unter optimalen Bedingungen. Nachteil der Per-protocol-Analyse ist, dass das Randomisierungsschema durchbrochen wird. Damit ist die Strukturgleichheit der zu vergleichenden Gruppen nicht mehr gegeben und es besteht die Möglichkeit systematisch verzerrter Ergebnisse. Per-protocol-Analysen werden vor allem in Interventionsstudien der Phasen I und II eingesetzt, um das therapeutische Potenzial und optimale Dosierungen für eine Behandlungsform zu finden. Das Gegenteil einer Per-Protocol-Analyse ist die *Intention-to-treat-Analyse*.

Placebo

In einer pragmatischen Definition ein meist zu Studienzwecken eingesetztes „Schein-Medikament“ ohne pharmakologisch aktive Substanz. Das Placebo darf hinsichtlich seiner äußeren Eigenschaften nicht von der aktiven Behandlung (dem Verum-Präparat) unterscheidbar sein, wenn es seinen Zweck erfüllen soll.

Positiver prädiktiver Wert

Anteil der Personen mit positivem Testergebnis, bei denen die gesuchte Krankheit auch tatsächlich vorliegt. Dieser Wert hängt von der *Prävalenz* der Erkrankung in der untersuchten Gruppe ab.

Power (statistische Trennschärfe)

Die Fähigkeit einer Studie, einen tatsächlich vorhandenen Unterschied statistisch signifikant (s. *statistische Signifikanz*) nachzuweisen und die *Nullhypothese* zu verwerfen, wenn sie

tatsächlich falsch ist. Der Nachweis bezieht sich auf a priori festgelegte Unterschiede in den Endpunkten (*Outcomes*) von Therapie- und Kontrollgruppe. Da die Power u.a. entscheidend vom Stichprobenumfang abhängt, kann der allgemein übliche Wert von 80% nur durch eine ausreichend große Stichprobe sichergestellt werden.

Prädiktion

Vorhersage, dass mit einer beschreibbaren Wahrscheinlichkeit ein Zustand oder ein Ereignis eintreten wird.

Prädiktoren/Risikofaktoren

Zustände, Merkmale oder Ereignisse, deren Vorhandensein oder Ausprägung die Vorhersage erlauben, dass ein bestimmter Zustand oder ein bestimmtes Zielereignis eintreten wird. Beispiele für Prädiktoren im Bereich Medizin sind genetische Merkmale, die zum Auftreten bestimmter Erkrankungen prädisponieren oder Risikofaktoren wie Diabetes für das Auftreten eines Herzinfarkts. Der Begriff Risikomarker wird oft verwendet, um eine Assoziation ohne gesicherte Kausalität auszudrücken, während der Begriff Risikofaktor häufig eine Kausalität impliziert. Die Nomenklatur ist hier jedoch nicht einheitlich. Wie gut sich *Outcomes* aus Prädiktoren voraussagen lassen, wird in der Epidemiologie mithilfe von Regressionsanalysen ermittelt.

Prävalenz

Die Prävalenz beschreibt den Anteil Erkrankter an der Gesamtzahl einer definierten Population zu einem bestimmten Zeitpunkt. (s. auch *Inzidenz*)

Primärliteratur

Publikationen, in denen Wissenschaftler ihre eigenen bzw. die Ergebnisse ihrer Forschungsgruppe (Primärforschung) veröffentlichen. (s. *Sekundärliteratur*)

Publikations-Bias (publication bias)

Systematischer Fehler (*Bias*) aufgrund einer selektiven Publikationspraxis, bei der Studien mit positiven und signifikanten Ergebnissen eine größere Chance haben, publiziert zu werden als Studien mit negativen und nicht-signifikanten Resultaten. Ein *systematischer Review* oder eine *Meta-Analyse*, die sich ausschließlich auf publizierte Studien stützt, läuft Gefahr, den Effekt der untersuchten Intervention zu überschätzen.

Randomisierte kontrollierte Studie (RCT)

Eine experimentelle Studie, bei der die Patienten nach einem Zufallsverfahren (mit *verdeckter Zuordnung*) auf die Therapie- bzw. die Kontrollgruppe verteilt (*Randomisierung*) und auf das Auftreten der festgelegten Endpunkte in den einzelnen Gruppen nachbeobachtet werden.

Randomisierung

Verfahren, das eine zufällige Verteilung der Patienten auf eine Therapie- und eine Kontrollgruppe bewirkt (s.a. randomisierte kontrollierte Studie). Dies kann durch (computergenerierte) Zufallszahlen oder andere Mechanismen erreicht werden. Damit soll sichergestellt werden, dass alle Teilnehmer die gleiche Chance haben, der einen oder anderen Gruppe zugeordnet zu werden und es wahrscheinlich ist, dass sich (bei ausreichender Studiengröße) bekannte wie unbekannte Risiko- und Prognosefaktoren ausgeglichen auf die beiden Gruppen verteilen. Wenn sich zwischen den beiden Gruppen in den Endpunkten ein Unterschied zeigt, kann dieser tatsächlich der experimentellen Intervention zugeordnet werden.

Referenzstandard (Goldstandard)

Bei der Erfassung der Testgenauigkeit (accuracy) von diagnostischen Tests gilt als Referenzstandard das Verfahren, das die derzeit beste zur Verfügung stehende Methode beschreibt und an dem neue bzw. alternative Methoden gemessen werden.

Relative Risiko-Reduktion (RRR)

Die relative Senkung der Rate an ungünstigen Ereignissen in der experimentellen Gruppe (E) einer Studie im Vergleich zur Kontrollgruppe. Sie wird wie folgt berechnet:

$$RRR = \frac{(K - E)}{K}$$

(s.a. absolute Risiko-Reduktion). Beispiel: Das Risiko für eine gastro-intestinale Blutung auf einer Intensivstation beträgt ohne Behandlung (Kontrollgruppe) 10% oder 0.10, bei Prophylaxe mit H₂-Blockern (E) 7% oder 0.07: Die RRR beträgt = 0.3 oder 30%.

Relatives Risiko (RR)

Effektmaß für dichotome Variablen. Das Relative Risiko in einer Therapiestudie bezeichnet das Verhältnis zwischen dem Risiko in der experimentellen Gruppe und dem Risiko in der Kontrollgruppe. Ein Relatives Risiko von 1 bedeutet, dass zwischen den Vergleichsgruppen kein Unterschied besteht (s. a. Absolute Risikoreduktion, Relative Risiko-Reduktion und Odds-Ratio). Bei ungünstigen Ereignissen zeigt ein RR <1, dass die experimentelle Intervention wirksam ist, um das Auftreten von ungünstigen Ereignissen zu senken.

Responder

1. Personen, die auf eine spezifische Behandlung (z.B. Medikamente) in der erwünschten Weise reagieren, z.B. Anteil von Patienten mit diabetischer Proteinurie, die auf ACE-Hemmer mit dem erwarteten Rückgang der Proteinurie reagiert (vergl. auch Therapieversager).
2. Personen einer Studienpopulation, die sich aktiv an der Studie, z.B. durch Rücksenden eines ausgefüllten Fragebogens, beteiligen.

Response-Rate

1. Anteil von Patienten, die auf eine Behandlung mit dem angestrebten Ergebnis reagieren (z.B. Rückgang der Tumorgroße).

2. Anteil der Studienpopulation, der sich aktiv an der Studie, z.B. einer Umfrage, beteiligt. Schicken beispielsweise die Hälfte der angeschriebenen Personen in einem Survey einen Fragebogen zurück, beträgt die Responserate 50%. Die Responserate sagt nichts über die Gültigkeit der Antworten aus – z.B. ob der Fragebogen korrekt ausgefüllt wurde. Je höher die Responserate in einer Studie ist, desto valider sind die Ergebnisse zu bewerten. Bei niedrigen Responseraten besteht immer die Gefahr von systematisch verzerrten Ergebnissen, da Charakteristika der nicht-antwortenden Studienpopulation möglicherweise in Beziehung zur Untersuchungsgröße der Studie stehen.

Risiko (Rate, Ereignisrate)

Die Wahrscheinlichkeit von Personen in einer Gruppe, bei denen ein bestimmter Endpunkt auftritt. Wenn z.B. in einer Gruppe von 100 Personen 30 einen bestimmten Endpunkt entwickeln (und bei 70 Personen das Ereignis nicht auftritt), ist das Risiko (oder die Ereignisrate) oder 0.3 oder 30%. (s. auch Odds)

Risikomarker (Risikoindikator)

Merkmal, welches auf ein erhöhtes Risiko für eine Krankheit oder ein negatives Outcome hinweist, ohne selbst kausal an der Entstehung beteiligt zu sein bzw. den Krankheitsverlauf zu beeinflussen. Die Beseitigung eines Risikomarkers verringert nicht das Krankheitsrisiko oder das Risiko eines ungünstigen Verlaufs. Ein typisches Beispiel für Risikomarker sind HLA-Antigene wie z.B. HLA-B-27 für die Entwicklung eines Morbus Bechterew.

Risikofaktor

s. Prädiktoren

Risikostratifikation

Aufteilung („Schichtung“) einer Studienpopulation in mehrere Untergruppen (Strata) in Abhängigkeit vom Risikostatus der Probanden. Die Risikostratifikation wird verwendet, um die Verzerrung von Studienergebnissen durch Confounder (*Confounding*) zu vermeiden. Z.B. ist bei vielen Krebserkrankungen ein fortgeschrittenes Krankheitsstadium ein bekannter Risikofaktor für ungünstige Behandlungsergebnisse (= schlechte Prognose). Daher werden die Auswertungen in Therapiestudien nach Krankheitsstadium stratifiziert, was eine Verzerrung der Gesamtergebnisse durch den Einschluss von Probanden mit unterschiedlichen Prognosen vermeidet und eine präzise Aussage zur Wirksamkeit innerhalb der jeweiligen Risikogruppe sicherstellt. (s. auch *Stratifizierung*)

Score

Punktwert, der aus der Erfassung einer Anzahl von Einzelinformationen in einem definierten Erfassungssystem nach einem festgelegten Algorithmus gebildet wird. Ein bekanntes Beispiel ist der APGAR-Score zur schnellen und systematischen Erfassung des Vitalzustands eines Neugeborenen nach der Geburt. Für Atmung, Herztätigkeit, Hautfarbe, Muskeltonus und Reflexe werden je 0 bis 2 Punkte vergeben und bilden aufsummiert den APGAR-Score. Ein APGAR-Score von 6 oder weniger zeigt einen behandlungsbedürftigen Zustand des Neugeborenen an.

Sekundärliteratur

Publikationen, in denen über die Werke anderer berichtet wird. Die Literaturübersicht, zu der auch systematische Übersichtsarbeiten (*systematische Reviews*) und Meta-Analysen gehören, ist eine Sonderform der Sekundärliteratur. Literaturübersichten fassen die wissenschaftlichen Erkenntnisse aus mehreren Einzelveröffentlichungen (*Primärliteratur*) zusammen.

Sensitivität

Anteil der Test-positiven Personen unter allen Erkrankten einer Stichprobe, d.h. die Wahrscheinlichkeit, mit einem diagnostischen Test die Kranken auch als krank zu identifizieren. Eine hohe Sensitivität wird angestrebt, wenn eine Erkrankung mit hoher Sicherheit ausgeschlossen werden soll. (s. auch *Spezifität*)

Sensitivitätsanalyse

Analyseverfahren, das feststellt, wie stabil sich eine Modellrechnung oder eine Meta-Analyse gegenüber Schwankungen der Eingangsparameter oder Veränderungen in der Berechnungsmethodik verhält. In gesundheitsökonomischen Modellrechnungen werden in einer Sensitivitätsanalyse mögliche Unterschiede in den Eingangsparametern (leichter versus schwerer Krankheitsschweregrad, gute versus schlechte Compliance, großer versus kleiner Behandlungseffekt) in ihren Auswirkungen auf die Endergebnisse untersucht und beschrieben.

Spezifität

Anteil der Test-negativen Personen unter allen Nicht-Erkrankten einer Stichprobe, d.h. die Wahrscheinlichkeit, mit einem diagnostischen Test Nicht-Erkrankte korrekt zu identifizieren. Eine hohe Spezifität wird angestrebt, wenn eine Erkrankung mit großer Sicherheit bestätigt werden soll.

Standardabweichung

Maß für die Streuung von Messwerten um den Durchschnittswert.

Statistische Signifikanz

Ein statistisch signifikantes Ergebnis einer Studie ist ein Ergebnis, das gegen die *Nullhypothese* spricht. Die Aussage basiert auf einem statistischen Test, der zur Prüfung einer vorab festgelegten Hypothese mit vorab festgelegter Irrtumswahrscheinlichkeit durchgeführt wird. Statistische Signifikanz darf nicht

mit klinischer Relevanz gleichgesetzt werden (s. auch *p-Wert*).

Stichprobe

Teilmenge der Gesamtpopulation, über die mit einer Studie eine Aussage gemacht werden soll. In der Stichprobe werden Daten erhoben, anhand derer Schlussfolgerungen für die gesamte Population gezogen werden (z.B. sollen von einer Studie über 200 Diabetiker mit Nierenbeteiligung Rückschlüsse gezogen werden auf alle Diabetiker mit Nierenbeteiligung). Um valide Aussagen für die gesamte Population zu erhalten, muss die Stichprobe 1. ausreichend groß sein, um den *Stichprobenfehler* klein zu halten und 2. repräsentativ für die gesamte Population sein, damit sie Strukturgleichheit mit der Population aufweist, für die die Aussage gemacht werden soll. Idealerweise wird dies durch eine zufällig gezogene Stichprobe erreicht.

Stichprobenfehler

Unsicherheit der Schätzung in einer *Stichprobe*. Jedes Merkmal einer Stichprobe liegt in der Gesamtpopulation mit einer „wahren Häufigkeit“ oder in einer „wahren Größe“ vor. Da die Stichprobe nur einen Teil der Population umfasst, ist damit zu rechnen, dass die Schätzung von Häufigkeit oder Größe des Merkmals in der Stichprobe von der Häufigkeit oder Größe in der Gesamtpopulation abweicht. Diese Abweichung wird als Stichprobenfehler bezeichnet. Der Stichprobenfehler ist umso größer, je kleiner die Stichprobe ist und je seltener das zu messende Merkmal in der Population vorkommt. Das Ausmaß des Stichprobenfehlers lässt sich durch Angabe des *Konfidenzintervalls* um den Punktschätzer beschreiben.

Stratifizierung

Aufteilung einer Stichprobe oder Untersuchungspopulation in verschiedene Schichten (*Strata*) in Abhängigkeit von der Ausprägung eines interessierenden Merkmals. Häufig

verwendete Stratifizierungsmerkmale sind Alter, Krankheits-Schweregrad, Geschlecht oder Ethnizität. Mit stratifizierten Ergebnisauswertungen kann man Ergebnisse für interessierende Subgruppen erhalten oder *Confounding* kontrollieren.

Studiendesign

Struktur einer problemgeleiteten, systematischen Beobachtung zur Beantwortung einer spezifischen Fragestellung (Studie). Das Studiendesign hängt von der Fragestellung ab und z.B. auch davon, ob es sich um ein qualitatives oder ein quantitatives Problem handelt. Bei **qualitativen Forschungsfragen** kommen unterschiedliche Methoden zum Einsatz: z.B. qualitative Interviews, Fokusgruppen; Literatur- oder Quellenanalysen. Bei **quantitativen Fragestellungen** (z.B. Häufigkeiten von Erkrankungen, Zusammenhang von Exposition und *Outcome*, Wirksamkeit einer Therapie) kommen Methoden der klinischen Epidemiologie oder der empirischen Sozialforschung zum Einsatz. Bei Beobachtungsstudien (z.B. *Kohortenstudien*; *Fall-Kontroll-Studien*; Querschnittsstudien oder ökologischen Studien) beschränkt sich die Tätigkeit der Wissenschaftler auf das Erheben und Auswerten von Daten. Bei Interventionsstudien (z.B. *randomisierten kontrollierten Studien*) wird von den Forschern eine Maßnahme (z.B. Behandlung, Aufklärungskampagne, Versorgungsform) gezielt eingesetzt, um in der Studienpopulation bestimmte *Outcomes* zu beeinflussen.

Surrogatendpunkte (intermediäre Endpunkte)

Endpunkte, die selbst nicht von unmittelbarer Bedeutung für die Patienten sind, aber stellvertretend für wichtige Endpunkte stehen können (z.B. Blutdrucksenkung als Surrogatparameter für Vermeidung eines Schlaganfalls). Surrogatendpunkte sind oft physiologische oder biochemische Marker, die sich relativ schnell und einfach messen

lassen und denen eine Vorhersagefunktion für spätere klinische Ereignisse zugestanden wird. Voraussetzung für zuverlässige Aussagen über die Wirksamkeit einer Behandlung ist ein enger kausaler Zusammenhang zwischen Surrogatparameter und dem eigentlichen Endpunkt. Für die allermeisten Surrogatparameter ist ein kausaler Zusammenhang nicht ausreichend belegt, was den Nutzen von Surrogatparametern für die Vorhersagekraft von klinisch relevanten Endpunkten infrage stellt. So haben z.B. in einer Studie mit antiarrhythmisch wirkenden Medikamenten (CAST-Studie) die geprüften Antiarrhythmika gefürchtete Herzrhythmusstörungen (Surrogat) verhindert – mittelfristig traten aber, im Vergleich zu den Kontrollgruppen, vermehrt Todesfälle (klinisch relevanter Endpunkt) auf.

Systematischer Fehler

s. Bias

Systematischer Review (Synonym:

Systematische Übersicht)

Sekundärforschung, bei der zu einer klar formulierten Frage alle verfügbaren Primärstudien systematisch und nach expliziten Methoden identifiziert, ausgewählt und kritisch bewertet und die Ergebnisse extrahiert und deskriptiv oder mit statistischen Methoden quantitativ (*Meta-Analyse*) zusammengefasst werden. Nicht jeder systematische Review führt zu einer Meta-Analyse.

Umschläge, verschlossene (sealed envelopes)

Unzuverlässige Methode der verdeckten randomisierten Zuordnung der Probanden zu Interventions- und Kontrollgruppe in RCTs. Für jeden einzelnen Studienteilnehmer liegt die Gruppenzuordnung in einem verschlossenen, undurchsichtigen Umschlag vor. Erst nach Überprüfung der Einschlusskriterien und Einwilligung des Probanden in die Studienteilnahme wird ein Umschlag geöffnet und die Gruppenzuordnung abgelesen. Die

Methode ist sehr manipulationsanfällig, da die Reihenfolge, in der die Umschläge geöffnet werden, nicht zuverlässig kontrolliert werden kann. Die Umschlagsmethode sollte daher nur dann angewandt werden, wenn eine zentrale *Randomisierung* nicht praktikabel ist, z.B. bei Studien in Notfallsituationen.

Validität

s. Interne Validität, Externe Validität

Verblindung

Geheimhaltung der Gruppenzuordnung (Therapie oder Kontrolle) vor Patienten, Studienärzten, Pflegepersonal und Auswertern, die an einer Studie teilnehmen. Damit soll verhindert werden, dass durch das Wissen um die Gruppenzugehörigkeit die Therapieantwort der Patienten, das Verhalten der Ärzte oder die Bewertung der Ergebnisse beeinflusst wird. In einfach-blinden Studien wissen nur die Patienten nicht über ihre Zuordnung Bescheid, in doppel-blinden Studien bleibt die Zuordnung Patient und Arzt verborgen. Die Verblindung von Ärzten und Patienten ist nicht immer durchführbar (z.B. beim Vergleich von chirurgischen mit medikamentösen Verfahren), wobei eine Verblindung der Endpunkt-Auswerter in der Regel möglich ist. (s. auch *Bias*)

Verdeckte Zuordnung (concealment of allocation)

Methodisches Verfahren zum Schutz vor Selektions-Bias. Geheimhaltung der randomisierten Zuteilungsfolge zu Therapie- oder Kontrollgruppe bis zum Zeitpunkt des Studieneinschlusses und der Zuordnung des Patienten zu einer Studiengruppe (s. auch *Bias*).

Vortest-Wahrscheinlichkeit einer Erkrankung (pre-test probability)

Wahrscheinlichkeit einer Erkrankung vor der Durchführung eines Tests. Sie beruht im Allgemeinen auf der *Prävalenz* der Erkran-

kung in einem bestimmten Umfeld (z.B. Normalbevölkerung, Primär-, Sekundärversorgung, im Krankenhaus, in der eigenen Praxis). Sind diese Informationen nicht verfügbar, müssen sie gegebenenfalls geschätzt werden.

Wahrscheinlichkeitsverhältnis (Likelihood-Ratio = LR)

Das Verhältnis der Wahrscheinlichkeit, dass ein positives (oder negatives) Testergebnis bei einer Person mit der Erkrankung auftritt, zu der Wahrscheinlichkeit, dass dieses positive (oder negative) Testergebnis bei einer Person ohne diese Erkrankung auftritt. Die LR ermöglicht eine Aussage darüber, wie stark sich durch das Testergebnis die Wahrscheinlichkeit für oder gegen das Vorliegen einer Erkrankung ändert (*Nachttest-Wahrscheinlichkeit*).

Zufallszahlen (random numbers)

In der klinischen Forschung dienen Zufallszahlen zur Bildung von Stichproben oder zur zufälligen Zuteilung von Probanden zur Interventions- oder Kontrollgruppe (*Randomisierung*). Echte Zufallszahlen werden z.B. beim Münzwurf, Würfeln oder Roulette erzeugt. In der Praxis wird meist auf Pseudo-Zufallszahlen zurückgegriffen, d.h. auf scheinbar zufällige Zahlen, die nach einem festen, reproduzierbaren Verfahren erzeugt werden. Obwohl nicht wirklich zufällig, haben sie ähnliche statistische Eigenschaften wie echte Zufallszahlenfolgen. (Pseudo-)Zufallszahlen können gedruckten Tabellen entnommen oder von Computerprogrammen erzeugt werden.

VIII Anhang

Für Ärzte	513
Datenbankrecherche: Exemplarische Suche in Medline über PubMed – 513	
<i>Zusammengestellt von Monika Lelgemann, MSc, Universität Bremen, Stand 02/2007</i>	
✂ Toolbox Evidenzquellen für den Arzt – 541	
✂ Toolbox Materialien für Patienten zur Dokumentation des Erkrankungsverlaufes – 542	
Für Patienten	543
✂ Toolbox Hilfestellung bei der Einschätzung der Qualität von Informationen im Internet – 543	
✂ Toolbox Verlässliche Informationsangebote im Internet – 544	
✂ Toolbox Unterstützung des Arzt-Patienten-Gesprächs – 545	
✂ Toolbox Erfahrungsaustausch mit ebenfalls Betroffenen – 546	
<i>Alle Toolboxes zusammengestellt von Dr. Sylvia Sängler (www.azq.de), Stand 07/2006</i>	

Die Texte im Anhang können auch als PDF-Dateien von der Seite des Deutschen Ärzte-Verlags heruntergeladen werden: <http://www.aerzteverlag.de/>

Datenbankrecherche

Exemplarische Suche in Medline über PubMed

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>

Dargestellt wird eine orientierende Suche, anhand derer einige der technischen Funktionen vorgestellt werden.

➤ Frage:

Therapie des „Aufmerksamkeitsdefizit-Syndrom“ bei Erwachsenen mit Ritalin®? Welche validen Daten liegen vor? Gesucht werden primär randomisierte kontrollierte Studien oder deren Meta-Analysen.

In der oben gestellten Frage sind folgende Elemente einer suchtauglichen Frage enthalten:

patient: erwachsene ADS Patienten

intervention: Ritalin

limits: entsprechend der „Hierarchie der Evidenz“ nur bestimmte Publikations(Studien)arten

Grundlagen: Spezifizität vs. Sensitivität

Spezifizität (specificity) = Relevanzrate


Die gefundenen Treffer sind zu einem großen Teil relevant (high precision)
z.B. Suche nach Schlagwort kombiniert mit Dokumenttyp
dafür werden nicht alle relevanten gefunden (low recall).

Sensitivität (sensitivity) = Vollständigkeitsrate

Die gefundenen Treffer enthalten möglichst vollständig alle relevanten Treffer (high recall)
z.B. Schlagwort- und Volltextsuche kombiniert
dafür sind zusätzlich viele unbrauchbare Treffer enthalten (low precision).

Grundlagen:

Trunkierung und Boolesche Operatoren

Suchmöglichkeit	Beispiele	Suchergebnis	
Trunkierung (auch: Wildcard) Symbole u. a.: *, ? z. T. automatisch	review* Studie? an?ssthesia *schmerz	review, reviews Studie, Studien, Studiendesign, Studienfahrt, Studienrat, ... anaesthesia, anesthesia Kopfschmerz, Rückenschmerz, ...	
Phrasensuche Symbol meistens "..." Manchmal Voreinstellung	"back pain"	Dokumente, die genau diese Wortkombination enthalten	
Boole'sche Operatoren zur logischen Verknüpfung der Suchbegriffe			
AND	back AND pain "low back pain" AND drugs	Nur Dokumente, die beide Wörter enthalten	
OR	Rückenschmerz* OR Kreuzschmerz* Sedativum OR Beruhigungsmittel anaesthesia OR anesthesia	Dokumente, die eines der Wörter oder beide Wörter enthalten	
NOT	"back pain" NOT low	Dokumente, die das erste, jedoch nicht das zweite Wort enthalten	
Klammerung (Gruppierung von Boole'schen Suchformulierungen für größere Komplexität)	(Sedativum OR Beruhigungsmittel) AND Nebenwirkungen	Dokumente, die immer das letzte Wort, und eines der beiden ersten Wörter oder beide enthalten	



Grundlagen: Feldtypen

Beschreibungskategorien eines Dokumentes (z.B. Artikel, Leitlinie) z.B. Autor, Titel, Erscheinungsjahr, Quelle (u.a. Zeitschriftentitel) und wichtig für themenbezogene Suchen wie hier

Schlagwörter (auch: Controlled terms, Subject headings) – Inhaltliche Beschreibung des Dokumentes; wird „von Hand“ vergeben; wichtigster Schlagwortthesaurus in der Medizin: **MeSH** = Medical Subject Headings

Dokumenttyp (auch: Document type, Publication type) – Formale Beschreibung der Art des Dokumentes (z.B. Artikel, Leitlinien, Studie); wird „von Hand“ vergeben

Volltextsuche (auch: Freitext, Free text) – Suche in allen Feldern der bibliographischen Beschreibung

The screenshot shows the PubMed homepage in a web browser. The browser's address bar displays "http://www.ncbi.nlm.nih.gov/pubmed". The page header includes the NCBI logo, the PubMed logo, and the text "A service of the National Library of Medicine and the National Institutes of Health". Below the header is a navigation menu with tabs for "All Databases", "PubMed", "Nucleotide", "Protein", "Genome", "Structure", "OMIM", "PMC", "Journals", and "Books". A search bar is prominently displayed with the text "Search PubMed for" and "Go" and "Clear" buttons. Below the search bar are buttons for "Limits", "Preview/Index", "History", "Clipboard", and "Details".

Two callouts provide instructions:

- Startseite nach Aufruf von Medline über PubMed:** A speech bubble pointing to the search bar area.
- Hier klicken um in die MeSH Datenbank zu kommen:** A box pointing to the "MeSH Database" link in the left sidebar.

The left sidebar contains a list of links under various categories:

- About Entrez
- Text Version
- Entrez PubMed Overview Help | FAQ Tutorials New/Noteworthy E-Utilities
- PubMed Services Journals Database MeSH Database Single Citation Matcher Batch Citation Matcher Clinical Queries Special Queries LinkOut My NCBI
- Related Resources Order Documents NLM Mobile NLM Catalog NLM Gateway TOXNET Consumer Health

The main content area features a "My NCBI" section with the text: "Set up an automated PubMed update in less than 5 minutes." followed by three numbered steps: (1) Get a My NCBI account, (2) Save your search, and (3) Your PubMed updates can be e-mailed directly to you. Below this is a link to "My NCBI Help" and a paragraph of introductory text about PubMed's services.

The screenshot shows the MeSH (Medical Subject Headings) database interface. At the top, a navigation bar includes 'All Databases', 'PubMed', 'Nucleotide', 'Protein', 'Genome', 'Structure', 'OMIM', 'PMC', 'Journals', and 'Books'. The search bar is set to 'MeSH' and contains the text 'for Ritalin'. A callout bubble points to the search bar with the text: 'An diesem Symbol können Sie immer erkennen, in welcher Datenbank Sie gerade sind'. Below the search bar, there are buttons for 'Limits', 'Preview/Index', 'History', 'Clipboard', and 'Details'. The main content area features a paragraph: 'MeSH is the U.S. National Library of Medicine's controlled vocabulary used for indexing articles for MEDLINE/PubMed. MeSH terminology provides a consistent way to... information that may use different terminology for the same concepts.' A callout bubble points to the word 'Ritalin' in the search bar with the text: 'Hier als nächsten Schritt Ritalin eingeben'. Below this, there are three bullet points: 'Use the MeSH database to find Medical Subject Heading T...', 'Searching with the MeSH Database', and 'Combining MeSH Terms'. A 'MeSH database tutorials:' section lists three items, each with a 'Quick Tour' icon: 'Searching with the MeSH Database', 'Combining MeSH Terms', and 'Applying Subheadings and other features of the MeSH Database'. A large callout box at the bottom center contains the text: 'Startseite der MeSH Datenbank Unterdatenbank der zur Verfügung stehenden Schlagwörter (MeSH = Medical Subject Heading)'. The left sidebar contains various links such as 'About Entrez', 'Text Version', 'Entrez PubMed Overview', 'Help | FAQ', 'Tutorials', 'New/Noteworthy', 'E-Utilities', 'PubMed Services', 'Journals Database', 'MeSH Database', 'Single Citation Matcher', 'Batch Citation Matcher', 'Clinical Queries', 'Special Queries', 'LinkOut', 'My NCBI', 'Related Resources', 'Order Documents', and 'NLM Mobile'.

The screenshot shows the PubMed search interface. At the top, the NCBI logo and 'PubMed' branding are visible. The search bar contains 'Methylphenidate'[MeSH]. Below the search bar, there are tabs for 'Limits', 'Preview/Index', 'History', 'Clipboard', and 'Details'. The results summary shows 'All: 3681' and 'Review: 360'. The first three results are listed, each with a checkbox and a document icon. Annotations are present: a grey box at the top right states 'Ergebnis der Suche nach dem MeSH Begriff Methylphenidate in Medline 3681 Treffer'; a grey box on the right side of the first result says 'Suche nach ADS als nächster Schritt, um mit „AND“ kombinieren zu können'; and a grey box at the bottom center says 'Zurück in die MeSH Datenbank – hier klicken'. The left sidebar contains navigation links for 'Entrez PubMed', 'Services', and 'Special Queries'.

Ergebnis der Suche nach dem MeSH Begriff Methylphenidate in Medline 3681 Treffer

Search PubMed for "Methylphenidate"[MeSH]

Display Summary Show 20 Sort by Send to

All: 3681 Review: 360

Items 1 - 20 of 3681 Page 1 of 185 Next

1: [Reimherr FW, Williams ED, Strong RE, Mestas R, So...](#) [Related Articles, Links](#)

A double-blind, placebo-controlled, crossover study of methylphenidate in adults with ADHD with assessed dimensions of the disorder.
J Clin Psychiatry. 2007 Jan;68(1):93-101.
PMID: 17284136 [PubMed - indexed for MEDLINE]

2: [van Vulpen LF, Malingre MM, Lange R, Bartelink AK.](#) [Related Articles, Links](#)

[Sympathomimetic syndrome caused by autointoxication with methylphenidate]
Ned Tijdschr Geneeskd. 2006 Dec 2;150(48):2661-4. Dutch.
PMID: 17205945 [PubMed - indexed for MEDLINE]

3: [Knight M.](#) [Related Articles, Links](#)

Stimulant-drug therapy for attention-deficit disorder (with or without hyperactivity) and

Zurück in die MeSH Datenbank – hier klicken

© Deutscher Ärzte-Verlag: R. Kunz et al., Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis, 2. Aufl.

NCBI MeSH National Library of Medicine NLM

Entrez PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search MeSH for Attention Deficit Disorder Go Clear

Limits Preview/Index History Clipboard Details

About Entrez

Text Version

Entrez PubMed

Overview Help | FAQ Tutorial New/Noteworthy E-Utilities

PubMed Services

Journals Database MeSH Database Single Citation Matcher Batch Citation Matcher Clinical Queries LinkOut Cubby

Related Resources

Order Documents NLM Gateway TOXNET Consumer Health Clinical Alerts ClinicalTrials.gov PubMed Central

"Methylphenidate" [MeSH]

„Suchkiste“

Search PubMed Clear

Suggestions: [Attention deficit disorder](#); [Attention deficit disorders](#); [Deficit disorder, attention](#); [Disorder, attention deficit](#); [Deficit disorders, attention](#); [Disorders, attention deficit](#); [Sex differentiation disorder](#); [Depersonalization disorder](#); [Child development disorder](#); [Sex differentiation disorders](#); [more...](#)

Display Summary Show: 20 Send to Search Box with AND

- Build a search strategy using the [Send to Search Box](#) feature
- Select a database (e.g., PubMed) under the Links menu

1: [Attention Deficit Disorder with Hyperactivity](#)

A behavior disorder originating in childhood in which the child exhibits a pattern of inappropriate inattention, impulsivity, and hyperactivity. Inattention and hyperactivity-impulsivity, one or the other or both, are the most frequent in males than females. Onset is in childhood. Symptoms are persistent although a minority experience the full complement of symptoms. Year introduced: 1984

1. Oben ADS eingeben
2. Entsprechender Begriff wird angezeigt
3. *Methylphenidate* aus der „Suchkiste“ löschen mit „Clear“
4. Attention Deficit Disorder anklicken
5. Send to „Search Box“

NCBI MeSH and the National Institutes of Health

All Databases PubMed Nucleotide Protein Genome Structure

Search MeSH for Attention Deficit Disorder Go Clear

Limits Preview/Index History Clipboard Details

"Attention Deficit Disorder with Hyperactivity"[MeSH]

Search PubMed Clear

Suggestions: [Attention deficit disorder](#); [Attention deficit disorders](#); [Deficit disorder, attention](#); [Disorder, attention deficit](#); [Deficit disorders, attention](#); [Disorders, attention deficit](#); [Sex differentiation disorder](#); [Depersonalization disorder](#); [Sex differentiation disorders](#); [Child development disorder](#); [more...](#)

Display Full Show 20 Send to

All: 1

- If making selections (e.g., Subheadings, etc.), use the [Send to Search Box](#) feature to see PubMed records with those specifications.
- Select PubMed under the Links menu to retrieve all records for the MeSH Term.
- Select [NLM MeSH Browser](#) under the Links menu for additional information.

1: **Attention Deficit Disorder with Hyperactivity** Links

A behavior disorder originating in childhood in which the essential features are signs of developmentally inappropriate inattention, impulsivity, and hyperactivity. Although most

NCBI
 A service of the National Library of Medicine
 and the National Institutes of Health
 www.pubmed.gov

All Databases PubMed Nucleotide Protein Genome Structure

Search PubMed for "Attention Deficit Disorder with Hyperactivity"[MeSH] Go Clear

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Send to

All: 10523 Review: 1653

Items 1 - 20 of 10523 Page 1 of 527 Next

1: [Reimherr FW, Williams ED, Strong RE, Mestas R, Marchant BK.](#) Related Articles, Links
 A double-blind, placebo-controlled, crossover study of the efficacy of a long-acting, extended-release oral system of methylphenidate in adults with ADHD with a range of symptom dimensions of the disorder.
 J Clin Psychiatry. 2007 Jan;68(1):93-101.
 PMID: 17284136 [PubMed - indexed for MEDLINE]

2: [Matas M.](#)
 Approach to attention deficit disorder in adults.
 Can Fam Physician. 2006 Aug;52:961-4.
 PMID: 17273498 [PubMed - indexed for MEDLINE]

3: [McLennan JD.](#) Related Articles, Links
 Deciding on stimulant use for childhood ADHD: maintaining focus on the key questions.
 Can Fam Physician. 2006 Aug;52:940-1, 945-6. English, French. No abstract available.
 PMID: 17273490 [PubMed - indexed for MEDLINE]

4: [Michelson D, Read HA, Ruff DD, Witcher J, Zhang S, McCracken J.](#) Related Articles, Links
 CYP2D6 and clinical response to atomoxetine in children and adolescents with ADHD.
 J Am Acad Child Adolesc Psychiatry. 2007 Feb;46(2):242-51.

Internet

So sollte es aussehen,
 nachdem Sie die Suche
 mit „Go“ gestartet haben.

10 523 Treffer zu ADS

In die „History“ Darstellung
 wechseln, um zu
 kombinieren

NCBI PubMed
A service of the National Library of Medicine and the National Institutes of Health
www.pubmed.gov

My NCBI [Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search PubMed for ((#16) AND (#17)) Preview Go Clear

Limits Preview/Ind History Clipboard Details

- Search History will be lost after eight hours of inactivity.
- Search numbers must not be continuous; all searches are re-processed.
- To save search indefinitely, click query # and select Save My NCBI.
- To combine searches use #search, e.g., #2 AND #3 or click query # for more options.

Search	Most Recent Queries	Time	Result
#17	Search "Attention Deficit Disorder with Hyperactivity" [H]	08:40:36	10523
#16	Search "Methylphenidate" [MeSH]	08:40:14	3681

Clear History

1. „Clear“, um ADS aus dem Suchfenster zu löschen
2. Kombination der Einzelergebnisse
3. Dann „Go“ um Suche zu starten

http://www.nih.gov/

Kombination der Einzelergebnisse: immer noch 1590 Treffer; weitere Eingrenzung auf Erwachsene erforderlich

NCBI PubMed
A service of the National Library of Medicine and the National Institutes of Health
www.pubmed.gov

All Databases PubMed Nucleotide Protein Genome Structure

Search PubMed for ((#16)) AND (#17) Go Clear Save

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by

All: 1590 Review: 234

Items 1 - 20 of 1590 Page 1 of 80 Next

1: [Reimherr FW, Williams ED, Strong RE, Mestas R, Soni P, Marchant BK.](#) Related Articles, Links
 A double-blind, placebo-controlled, crossover study of osmotic release oral system methylphenidate in adults with ADHD with assessment of oppositional and emotional dimensions of the disorder.
 J Clin Psychiatry. 2007 Jan;68(1):93-101.
 PMID: 17284136 [PubMed - indexed for MEDLINE]

2: [Knight M.](#) Related Articles, Links
 Stimulant-drug therapy for attention-deficit disorder (with or without hyperactivity) and sudden cardiac death.
 Pediatrics. 2007 Jan;119(1):154-5. No abstract available.
 PMID: 17200281 [PubMed - indexed for MEDLINE]

3: [Sharp BW.](#) Related Articles, Links

Fertig Internet

Freitextsuche nach *adult**

NCBI PubMed
www.pubmed.gov

Search PubMed for **adult*** Go Clear Save Search

Display Summary Show 20 Sort by Send to

All: 3140400 Review: 140077

Items 1 - 20 of 3140400 Page 1 of 157020 Next

- 1: [Arslan M, Karadeniz AN, Aksu G, Gurveli M, Fayda M, Dogan AK, Akyuz F.](#)
Postoperative hypofractionated radiotherapy in glioblastoma multiforme.
J BUON. 2006 Jan-Mar;11(1):39-42.
PMID: 17318950 [PubMed - in process]
- 2: [Garman R, Gaudette G, Donahue LR, Rubin C, Judex S.](#)
Low-level accelerations applied in the absence of weight bearing can enhance trabecular bone formation.
J Orthop Res. 2007 Feb 22; [Epub ahead of print]
PMID: 17318899 [PubMed - as supplied by publisher]
- 3: [Silfversward CJ, Larsson S, Ohlsson C, Frost A, Nilsson O.](#)
Reduced cortical bone mass in mice with inactivation of interleukin-4 and interleukin-13.
[print]

* entspricht einem Trunkierungszeichen – es wird nach variablen Wortendungen gesucht

adult[All Fields] OR adult/acute[All Fields] OR
adult/adolescent[All Fields] OR adult/adult[All Fields] OR
adult/age[All Fields] OR adult/aged[All Fields] OR
adult/aging[All Fields] OR adult/child[All Fields] OR
adult/chronic[All Fields] OR adult/complications[All Fields] OR
adult/cryotop[All Fields] OR adult/developmental[All Fields] OR
adult/donor[All Fields] OR adult/eggs[All Fields] OR
adult/elderly[All Fields] OR adult/elite[All Fields] OR
adult/embryonic[All Fields] OR adult/etiology[All Fields] OR
adult/family[All Fields] OR adult/fetal[All Fields] OR
adult/flanged[All Fields] OR adult/geriatric[All Fields] OR
adult/gerontological[All Fields] OR adult/human[All Fields] OR
adult/immature[All Fields] OR adult/juvenile[All Fields] OR
adult/larvae[All Fields] OR adult/late[All Fields] OR adult/m2[All
Fields] OR adult/metabolism[All Fields] OR adult/neonatal[All
Fields] OR adult/newborn[All Fields] OR adult/nymph[All Fields]
OR adult/nymphs[All Fields] OR adult/old[All Fields] OR
adult/older[All Fields] OR adult/ops[All Fields] OR
adult/paediatic[All Fields] OR adult/parent[All Fields] OR
adult/pathology[All Fields] OR adult/pediatric[All Fields] OR
adult/physiopathology[All Fields] OR adult/psychology[All
Fields] OR adult/pubertal[All Fields] OR adult/radiotherapy[All
Fields] OR adult/sex[All Fields] OR adult/sibling[All Fields] OR
adult/somatic[All Fields] OR adult/student[All Fields] OR Fields]

Im
Hintergrund
sucht
MEDLINE bei
Eingabe von
adult* als
Freitextsuche
nach allen
dargestellten
Begriffen
(Auszug)

The screenshot shows the PubMed website interface. At the top, there is a navigation bar with "All Databases", "PubMed", "Nucleotide", "Protein", "Genome", "Structure", "OMIM", "PMC", "Journals", and "Books". The search bar contains "PubMed" and "for Adult*", with "Go", "Clear", and "Save Search" buttons. Below the search bar are tabs for "Limits", "Preview/Index", "History", "Clipboard", and "Details". The "History" tab is highlighted, and a callout box points to it with the text: "Wechseln in die „History“-Darstellung, um durch weitere Kombination einzugrenzen". The main content area displays search results for "All: 3140400" and "Review: 140077". The results list includes:

- 1: Arslan M, Karadeniz AN, Aksu G, Guveli ... da M, Dogan AK, Akyuz F.**
Postoperative hypofractionated radiotherapy in ... glioblastoma multiforme.
J BUON. 2006 Jan-Mar;11(1):39-42.
PMID: 17318950 [PubMed - in process]
- 2: Gaman R, Gaudette G, Donahue LR, Rubin C, Jud...**
Low-level accelerations applied in the absence of ... ing can enhance trabecular bone formation.
J Orthop Res. 2007 Feb 22; [Epub ahead of p...]
PMID: 17318899 [PubMed - as supplied by p...]
- 3: Silfversward CJ, Larsson S, Ohlsson C, Fros...**
Reduced cortical bone mass in mice with ...
J Orthop Res. 2007 Feb 22; [Epub ahead of print]
PMID: 17318894 [PubMed - as supplied by publisher]
- 4: Archambault JM, Jelinsky SA, Lake SP, Hill AA, Glaser DL, Soslowsky LJ.**
Rat supraspinatus tendon expresses cartilage markers with overuse.
J Orthop Res. 2007 Feb 22; [Epub ahead of print]
PMID: 17318892 [PubMed - as supplied by publisher]

The interface also shows a sidebar with "Entrez PubMed" services and a footer with "Fertig" and "Internet" icons.

Suchschritt #18 und Suchschritt #19 zur weiteren Eingrenzung kombinieren

The screenshot shows the PubMed interface with a search history table. A callout box points to the first two entries in the history table, indicating that steps #18 and #19 should be combined for further refinement.

Search History Instructions:

- Search History will be lost after eight hours of inactivity.
- Search numbers may not be continuous; all searches are represented.
- To save search indexes locally, click query # and select Save in My NCBI.
- To combine searches use #search, e.g., #2 AND #3 or click query # for more options.

Search	Most Recent Queries	Time	Result
#19	Search adult*	08:44:36	<u>3140400</u>
#18	Search ((#16)) AND (#17)	08:43:08	<u>1590</u>
#17	Search "Attention Deficit Disorder with Hyperactivity"[MeSH]	08:40:36	<u>10523</u>
#16	Search "Methylphenidate"[MeSH]	08:40:14	<u>3681</u>

Buttons: Clear History

Kombination mit „AND“

Kombination ergibt 307 Treffer; immer noch zu viele, um sich schnell zu orientieren

Weitere Eingrenzung über den Studientyp (Publikationstyp)

NCBI PubMed A service of the National Library of Medicine and the National Institutes of Health www.pubmed.gov

All Databases PubMed Nucleotide Protein Genome Struct

Search PubMed for ((#18)) AND (#19) Go

Limits Preview/Index History Clipboard

Display Summary Show 20 Send to

All: 307 Review: 76

Items 1 - 20 of 307 Page 1 of 16 Next

1: [Reimherr FW, Williams ED, Strong RE, Mestas R, Soni P, Marchant BK.](#) Related Articles, Links

A double-blind, placebo-controlled, crossover study of osmotic release oral system methylphenidate in adults with ADHD with assessment of oppositional and emotional dimensions of the disorder. J Clin Psychiatry. 2007 Jan;68(1):93-101. PMID: 17284136 [PubMed - indexed for MEDLINE]

2: [\[No authors listed\]](#) Related Articles, Links

ADHD update: new data on the risks of medication. Harv Ment Health Lett. 2006 Oct;23(4):3-5. No abstract available. PMID: 17183739 [PubMed - indexed for MEDLINE]

3: [Fallu A, Richard C, Prinzo R, Binder C.](#) Related Articles, Links

Does OROS-methylphenidate improve core symptoms and deficits in executive function?

The screenshot shows the PubMed website interface. At the top, there is a navigation bar with the NCBI logo and the text 'A service of the National Library of Medicine and the National Institutes of Health'. Below this is a search bar with the text 'Search PubMed for [(#5)) AND (#7)'. The 'Limits' menu is highlighted, and a callout box points to it with the text: 'Weitere Eingrenzung über den Studientyp (Publikationstyp) Mit der „Limits“-Funktion möglich; bei Anklicken öffnet sich ein Untermenü'. The 'Limits' menu includes options for 'Search by Author', 'Search by Journal', 'Full Text, Free Full Text, and Abstracts', 'Dates', 'Humans or Animals', 'Languages', and 'Subsets'.

The screenshot shows the NCBI search interface with the following filter sections:

- My NCBI** (Left sidebar):
 - Related Resources
 - Order Documents
 - NLM Mobile
 - NLM Catalog
 - NLM Gateway
 - TOXNET
 - Consumer Health
 - Clinical Alerts
 - ClinicalTrials.gov
 - PubMed Central
- Languages**:
 - English
 - French
 - German
 - Italian
 - Japanese
 - Russian
 - Spanish
 - More Languages**
 - Afrikaans
 - Albanian
- Subsets**:
 - Bioethics
 - Cancer
 - Complementary Medicine
 - History of Medicine
 - Space Life Sciences
 - Systematic Reviews
 - Toxicology
 - More Subsets**
 - MEDLINE
 - PubMed Central
- Type of Article**:
 - Clinical Trial
 - Editorial
 - Letter
 - Meta-Analysis
 - Practice Guideline
 - Randomized Controlled Trial
 - Review
 - More Publication Types**
 - Addresses
 - Bibliography
- Ages**:
 - All Infant: birth-23 months
 - All Child: 0-18 years
 - All Adult: 19+ years
 - Adult: 19-44 years
 - Middle Aged: 45-64 years
- Tag Terms**:
 - Default Tag:

Buttons: **GO**, **Clear All Limits**

Footer: [Write to the Help Desk](#), [NCBI](#) | [NLM](#) | [NIH](#), [Department of Health & Human Services](#), [Privacy Statement](#) | [Freedom of information Act](#) | [Disclaimer](#)

Internet

The screenshot shows the PubMed interface with the following elements:

- Header:** NCBI logo, PubMed logo, and text: "A service of the National Library of Medicine and the National Institutes of Health".
- Navigation:** "All Databases", "PubMed", "Nucleotide", "Protein", "Genome".
- Search Bar:** "Search PubMed" with a dropdown menu set to "PubMed" and a search query: "for ((#18)) AND (#19)".
- Buttons:** "Limits", "Preview/Index", "History", "Clipboard".
- Results Summary:** "Limits: Randomized Controlled Trial", "Display Summary", "Show 20", "All: 45", "Review: 1".
- Page Info:** "Items 1 - 20 of 45", "Page 1 of 3 Next".
- Search Results:**
 - 1:** Reimherr FW, Williams ED, Strong RE, Mestas R, Soni P, Marchant BK. [Related Articles, Links](#)
A double-blind, placebo-controlled, crossover study of osmotic release oral system methylphenidate in adults with ADHD with assessment of oppositional and emotional dimensions of the disorder.
J Clin Psychiatry. 2007 Jan;68(1):93-101.
 PMID: 17284136 [PubMed - indexed for MEDLINE]
 - 2:** Wilson HK, Cox DJ, Merkel RL, Moore M, Coghill D. [Related Articles, Links](#)
Effect of extended release stimulant-based medications on neuropsychological functioning among adolescents with Attention-Deficit/Hyperactivity Disorder.
Arch Clin Neuropsychol. 2006 Dec;21(8):797-807. Epub 2006 Oct 17.
 PMID: 17049803 [PubMed - indexed for MEDLINE]

Callout Box Text:

- Einschränkung ergibt 45 Treffer, immer noch zu viele ?
- Weitere Eingrenzung auf Meta-Analysen möglich
- Wieder Verwendung der „Limits“-Funktion

The screenshot shows a web browser window with the URL <http://www.ncbi.nlm.nih.gov>. The interface includes several filter sections:

- Languages:** German, Italian, Japanese, Russian, Spanish, Afrikaans, Albanian.
- More Languages:** (dropdown menu)
- Type of Article:** Clinical Trial, Editorial, Letter, Meta-Analysis, Practice Guideline, Randomized Controlled Trial, Review.
- More Publication Types:** Addresses, Bibliography.
- Topics:** Dental journals, Nursing journals, AIDS, Bioethics, Cancer, Complementary Medicine, History of Medicine.
- Ages:** All Infant: birth-23 months, All Child: 0-18 years, All Adult: 19+ years, Newborn: birth-1 month, Infant: 1-23 months, Preschool Child: 2-5 years, Child: 6-12 years, Adolescent: 13-18 years, Adult: 19-44 years, Middle Aged: 45-64 years.
- Tag Terms:** Default Tag: All Fields.

A large grey arrow points from the 'Meta-Analysis' checkbox in the 'Type of Article' section to the 'All Fields' dropdown in the 'Tag Terms' section. Below the filters are 'GO' and 'Clear All Limits' buttons. At the bottom, there is a footer with contact information and a date: 'Feb 21 2007 18:14:20'.

3 Treffer!!!

Zur Darstellung der Abstracts
öffnen Sie das Drop-down-Menü

www.pubmed.gov

All Databases PubMed Nucleotide Protein Gene

Search PubMed for ((#18)) AND (#19)

Limits Preview/Index History Clipboard Details

Limits: Meta-Analysis

Display Summary Show 20 Sort by Send to

All: 3 Review: 0

Items 1 - 3 of 3 One page.

1: Faraone SV, Spencer T, Aleardi M, Pagano C, Biederman J. Related Articles, Links
Meta-analysis of the efficacy of methylphenidate for treating adult attention-deficit/hyperactivity disorder.
 J Clin Psychopharmacol. 2004 Feb;24(1):24-9.
 PMID: 14709943 [PubMed - indexed for MEDLINE]

2: Schachter HM, Pham B, King J, Langford S, Moher D. Related Articles, Links
How efficacious and safe is short-acting methylphenidate for the treatment of attention-deficit disorder in children and adolescents? A meta-analysis.
 CMAJ. 2001 Nov 27;165(11):1475-88.
 PMID: 11762571 [PubMed - indexed for MEDLINE]

3: Jadad AR, Boyle M, Cunningham C, Kim M, Schachar R. Related Articles, Links
Treatment of attention-deficit/hyperactivity disorder.
 Evid Rep Technol Assess (Summ). 1999 Nov;(11):i-viii, 1-341.
 PMID: 10790990 [PubMed - indexed for MEDLINE]

Internet

The screenshot shows the Entrez PubMed interface. A search for 'pubmed' has been performed, resulting in 3 items. The interface includes a navigation menu on the left, a search bar, and a list of search results. Two callout boxes provide additional information:

- Abstract als Darstellungsform öffnen:** A callout box pointing to the 'Abstract' option in the dropdown menu, indicating that clicking this option opens the full text of the article.
- Dieses Symbol zeigt an, dass der Volltext der Arbeit frei zur Verfügung steht (Klicken und es öffnet sich der entsprechende Link):** A callout box pointing to a document icon next to the first search result, indicating that the full text of the article is available for free.

The search results shown are:

- 1: Fagnano M, Pagano C, Biederman J. Efficacy of methylphenidate in the treatment of attention deficit hyperactivity disorder. J Clin Psychopharmacol. 2004 Feb;24(1):24-9. PMID: 14709943 [PubMed - indexed for MEDLINE]**
- 2: Schachter HM, Pham P, et al. How efficacious and safe is short-acting methylphenidate in the treatment of attention deficit hyperactivity disorder in children and adolescents? A meta-analysis. CMAJ. 2001 Nov 27;165(11):1475-88. PMID: 11762571 [PubMed - indexed for MEDLINE]**
- 3: Jadad AR, Boyle M, Cunningham C, Kim M, Schachar R.**

The screenshot shows a web browser window with the address bar displaying http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Search&utool=pubmed_Abstract&term=%22Pagano+C%22%5BAuthor%22. The browser interface includes a menu bar (Datei, Bearbeiten, Ansicht, Adressen, Wechsellisten, Links) and a toolbar. The main content area is titled "Limits: Meta-Analysis" and shows search filters: "Display: Abstract", "Show: 20", "Sort by:", and "Send to:". Below the filters, it indicates "All: 3", "Review: 0", and "Items 1 - 3 of 3". A single search result is displayed, titled "1: J Clin Psychopharmacol. 2004 Feb;24(1):24-9." with a link to "Related Articles, Links". The authors listed are "Faraone SV, Spencer T, Aleardi M, Pagano C, Biederman J." The abstract text reads: "This article reviews the efficacy of methylphenidate (MPH) for adult attention-deficit/hyperactivity disorder (ADHD). A literature search identified double-blind placebo-controlled MPH treatment studies of ADHD adults. Meta-analysis estimated the pooled effect size for MPH treatment and tested for publication bias. Meta-analysis regression assessed the influence of study design features on medication effects. Six trials met criteria and were included in this meta-analysis. These studies included a total of 140 MPH-treated ADHD adults and 113 placebo-treated ADHD adults. The mean effect size of 0.9 was statistically significant and there was no evidence of publication bias. Larger MPH effect sizes were associated with physician ratings of outcome and use of higher doses. When treatment is optimized to high doses, the effect size for MPH in adults was 1.3. We found strong support for the assertion that MPH is efficacious for treating adult ADHD. Because the degree of efficacy of MPH in treating ADHD adults is similar to what has been reported from meta-analyses of the child and adolescent literature, our work provides further assurance to clinicians that the diagnosis of ADHD can be validly applied in adulthood."

Zusammenfassende Darstellung der Suchschritte über die „History“ – eignet sich auch zur Dokumentation der Suche

NCBI PubMed
 A service of the National Library of Medicine and the National Institutes of Health
 www.pubmed.gov

All Databases PubMed Nucleotide Protein Gene

Search PubMed for [(#18)) AND (#19) Preview

Limits Preview/Index **History** Clipboard Details

Limits: Meta-Analysis

- Search History will be lost after eight hours of inactivity.
- Search numbers may not be continuous; all searches are represented.
- To save search indefinitely, click query # and select Save in My NCBI.
- To combine searches use #search, e.g., #2 AND #3 or click query # for more options.

Search	Most Recent Queries	Time	Result
#22	Search ((#18)) AND (#19) Limits: Meta-Analysis	08:49:31	3
#21	Search ((#18)) AND (#19) Limits: Randomized Controlled Trial	08:48:39	45
#20	Search ((#18)) AND (#19)	08:46:52	307
#19	Search adult*	08:44:36	3140400
#18	Search ((#16)) AND (#17)	08:43:08	1590
#17	Search "Attention Deficit Disorder with Hyperactivity"[MeSH]	08:40:36	10523
#16	Search "Methylphenidate"[MeSH]	08:40:14	3681

Clear History

Bei Fragen oder Unklarheiten verwenden Sie die gut verständliche „Hilfefunktion“ von PubMed

PubMed Help
The National Library of Medicine

Home | About Us | Contact Us | Feedback | Privacy Policy | Terms of Use | PubMed Help Contents | Bookshelf Help | NCBI | Browse More Books

Navigation

[About this book](#)

[PubMed Help](#)

[PubMed Quick Start](#)

[Search Field Descriptions and Tags](#)

[Searching PubMed](#)

[Understanding Your Search Results](#)

[Displaying and Sorting Your Search Results](#)

[Changing How Search Results are Displayed](#)

[Finding Related Resources for a Citation](#)

PubMed Help

Created: August 01, 2004

Updated: December 11, 2006

PubMed Help

PubMed lets you search more than 16 million bibliographic citations and abstracts in the fields of medicine, nursing, dentistry, veterinary medicine, the health care system, and preclinical sciences. It provides access to MEDLINE® and to articles in selected life sciences journals not included in MEDLINE.

PubMed is a free resource. It was developed and maintained by the National Center for Biotechnology Information (NCBI) at the U.S. National Library of Medicine (NLM) located at the National Institutes of Health.

For a brief overview of searching PubMed, see the [PubMed Quick Start](#) section.

This PubMed Help book was prepared by the [NCBI](#).

Nützliche Literaturhinweise:

Motschall E, Falck-Ytter Y: Searching the MEDLINE Literature Database through PubMed: A Short Guide. *Onkologie*, 2005; 28 (10) : 517-522

- **Literatursuche mit der medizinischen Datenbank MEDLINE**
EbM-Splitter Ausgabe 06/2004
- **Literatursuche in PubMed: Medical Subject Headings (MeSH)**
EbM-Splitter Ausgabe 08/2003
- **Medline-Recherche: PubMed Clinical Queries**
EbM-Splitter Ausgabe 10/2003

Alle unter:

http://www.ebm-netzwerk.de/grundlagen/grundlagen/splitter_zahnmedizin

Ansonsten hilft nur üben, üben, üben...

Hinweise auf Kurse, in denen Sie die Techniken der Datenbankrecherche erlernen finden Sie auf der Internetseite des Deutschen Netzwerkes für Evidenzbasierte Medizin: www.ebm-netzwerk.de

FÜR ÄRZTE

Toolbox Evidenzquellen für den Arzt

Um die Flut an vorhandenen Links einzugrenzen, haben wir uns darauf beschränkt, in erster Linie Sammlungen entsprechender Quellen anzugeben und nicht die Quellen einzeln anzuführen. Diese Zusammenstellung erhebt keinen Anspruch auf Vollständigkeit.

Recherche in:	
Literaturdatenbanken	<ul style="list-style-type: none"> • MEDLINE: http://www.pubmed.org • Das Deutsche Institut für medizinische Dokumentation und Information bietet unter http://www.dimdi.de/static/de/db/index.htm die Möglichkeit der Recherche in 70 verschiedenen (teils kostenpflichtigen) Datenbanken an. <p>Anmerkung: Formale Suchbegriffe: guideline(s), practice guideline(s), clinical practice guideline(s), recommendation(s), consensus/position statement(s), standard(s), Empfehlung(en), Richtlinie(n), Leitlinie(n) + Erkrankungsbegriff</p>
Leitliniendatenbanken, Leitliniensammlungen	<ul style="list-style-type: none"> • Leitlinien der Fachgesellschaften der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften: http://leitlinien.net • Datenbank deutschsprachiger und fremdsprachiger medizinischer und Pflege-Leitlinien zu verschiedenen Erkrankungsthemen: http://www.leitlinien.de • Leitliniendatenbank des Guidelines International Network: http://www.g-i-n.net Die Datenbank ist die weltweit umfangreichste Leitliniendatenbank, die derzeit ca. 4 000 Dokumente enthält (Stand: 02/2007) • Leitliniendatenbank des Nationalen Guideline Clearinghouse: http://www.guideline.gov • Unter der Adresse http://www.akdae.de/35/index.html können die Therapieempfehlungen der Arzneimittelkommission der deutschen Ärzteschaft abgerufen werden. • Nationale Versorgungsleitlinien unter Trägerschaft von Bundesärztekammer, Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften und Kassenärztlicher Bundesvereinigung: http://www.versorgungsleitlinien.de
Studien	<ul style="list-style-type: none"> • Cochrane Library (mit einführenden Erklärungen des Deutschen Cochrane Zentrums): http://www.cochrane.de/de/clibintro.htm • Quellen für Studien in der Linksammlung des DNEbM e.V.: http://www.ebm-netzwerk.de/literatur_und_links/linksammlung#11 • Arbeitsmaterialien zu EbM des DNEbM e.V.: http://www.ebm-netzwerk.de/grundlagen/grundlagen/arbeitsmaterialien • Das in der Schweiz ansässige Horton Zentrum für praxisorientierte Forschung und Wissenstransfer veröffentlicht unter http://www.evimed.ch derzeit bereits 140 Zusammenfassungen klinischer Studien zu verschiedenen Fachgebieten.
Wissenschaftliche Arbeiten im Internet finden	<ul style="list-style-type: none"> • Die Adresse http://www.scholar.google.com ist ein Serviceangebot der Suchmaschine Google. Es ermöglicht die Suche nach wissenschaftlicher Literatur. • Scirus (http://www.scirus.com) ist eine Spezialsuchmaschine für wissenschaftliche, technische und medizinische Inhalte. Ihre Beschränkung auf Universitäten, ausgewählte Server und wissenschaftliche Webseiten sorgt für eine übersichtliche Ergebnisliste frei von Ballast.
Internetsuchmaschinen	<ul style="list-style-type: none"> • Die gebräuchlichste Suchmaschine ist Google: http://www.google.de • Themenorientierte Suchmaschine der Health on the Net-Foundation zu geprüften medizinischen Angeboten: http://www.hon.ch/HONcode/Hunt • Suchmaschinen bieten zwar eine hohe Treffermenge, liefern aber viele unbrauchbare Treffer. Sie sind daher sinnvoll als Zusatzrecherche. Lesen Sie die entsprechenden Suchstrategien im Hilfetext nach.
<i>Zusammengestellt von Dr. Sylvia Sängler (www.azq.de), Stand: 07/2006</i>	

FÜR ÄRZTE



Toolbox Materialien für Patienten zur Dokumentation des Erkrankungsverlaufes

Unter den nachfolgenden Adressen finden Sie verschiedene Dokumentationshilfen, die an Patienten weitergegeben werden können, damit diese ihren Erkrankungsverlauf selbstständig dokumentieren können. Raten Sie Ihren Patienten, die entsprechenden Unterlagen auch und gerade im Falle eines Arztwechsels weiter zu nutzen.

Thema	Quelle
Schmerz	Schmerztagebücher/Schmerzfragebögen http://www.schmerzliga.de/aktionen/pain_check.htm http://www.krebsinformation.de/ksid/Materialien/Schmerztagebuch.pdf http://www.medizin.uni-koeln.de/stan/Schmerzmanual/TU/tagebuch.html http://www.schmerzambulanz.humanmedizin-goettingen.de/pdf/tagebuch_kopfschmerz.pdf http://www.tumorschmerz.de/Patiententagebuch.doc
Tumoren	Tumornachsorgepass http://www.thueringische-krebsgesellschaft.de/qualitaetsicherung/tumornachsorgepass.shtml oder Nachfrage bei der Deutschen Krebsgesellschaft: http://www.krebsgesellschaft.de bzw. der Deutschen Krebshilfe http://www.krebshilfe.de
Diabetes	Diabetespass Anbieter: Deutscher Hausärzteverband: http://www.ifap.de/bda-manuale/diabetes/komplikationen/diabetespass.html Patienten können sich an die Geschäftsstelle ihrer Krankenkasse wenden.
Erkrankungen der Lunge	Allergiepass Patienten können sich an die Geschäftsstelle ihrer Krankenkasse wenden. Einen Editor zum Erstellen personalisierter Allergiepässe kann als Freeware geladen werden unter: http://www.ivdk.gwdg.de/dkg/dok/AllergiepassEditor.zip Asthmatagebuch Patienten können sich an die Geschäftsstelle ihrer Krankenkasse wenden. Anbieter: Universität Zürich http://www.asthmainfo.ch/bilder_illu/asthmatagebuch_deutsch.pdf Beim Deutschen Allergie- und Asthmabund anfordern: http://www.daab.de/
Bluthochdruck	Blutdruckpass Patienten können sich an die Geschäftsstelle ihrer Krankenkasse oder Apotheken wenden. Online Bestellung unter: http://www.hypertonie-schulungsprogramm.de/
Allgemein	Patiententagebuch Anbieter: Verbraucherzentrale NRW, Patiententagebuch kann unter der Adresse: http://www.verbraucherzentrale-nrw.de/UNIQ115261186223308/link195213A.html zum Preis von 1,20 € heruntergeladen werden. Anbieter: Verbraucherzentrale Hessen, kann in der Ratgeberreihe bestellt werden unter: http://www.verbraucher.de/ratgeber/index.html zum Preis von 1,53 €
Zusammengestellt von Dr. Sylvia Sängler (www.azq.de), Stand: 07/2006	



Für Patienten



Toolbox
Hilfestellung bei der Einschätzung
der Qualität von Informationen
im Internet

Ein Service Ihres Arztes:

Liebe Patientin/Lieber Patient,
 wenn Sie sich über Ihre Erkrankung und alles, was damit zusammen hängt, informieren möchten, sollten Sie bedenken, dass nicht jede Information von guter Qualität ist. Das gilt ganz besonders für Informationen, die Sie im Internet finden.

Eine gute Patienteninformation enthält:

- ▲ Klare Angaben zu behandelten Themen und Zielgruppen
- ▲ Angaben von Literatur-Quellen und weiteren Hilfsangeboten
- ▲ Angaben zur Qualifikation des Autors
- ▲ Angaben zur Aktualität und Gültigkeit der Information
- ▲ Hinweise auf Ausgewogenheit und Unabhängigkeit
- ▲ Beschreibung von Unsicherheiten (Was ist noch nicht bewiesen?)
- ▲ Darstellung aller möglichen Behandlungsoptionen
- ▲ Erläuterung von Nutzen und Risiken der dargestellten Methoden
- ▲ Erklärung zu Auswirkungen der Behandlung auf das tägliche Leben
- ▲ Erklärung des natürlichen Krankheitsverlaufes

Wenn ein Anbieter von Informationen mit einer Qualitätsprüfung seines Internetangebots und/oder der darauf angebotenen Informationen wirbt, dann suchen Sie nach Hinweisen dazu:

- ▲ Nach welchen Kriterien geprüft wurde
- ▲ Ob man diese Kriterien nachlesen kann
- ▲ Wann die Qualitätsprüfung erfolgt ist
- ▲ Wer die Qualitätsprüfung vorgenommen hat
- ▲ Welches Ziel die Qualitätsprüfung hatte
- ▲ Wer diese Art Prüfung noch verwendet

Was Sie beim Stöbern im Internet beachten sollten

- ▲ Lesen Sie die Informationen immer kritisch
- ▲ Holen Sie immer mehrere Informationen ein
- ▲ Vergleichen Sie die Inhalte unterschiedlicher Informationen
- ▲ Sprechen Sie mit Ihrem Arzt über gefundene Informationen
- ▲ Hinterfragen Sie immer die Qualität der gefundenen Informationen
- ▲ Orientieren Sie sich an Qualitätskriterien für Patienteninformationen
- ▲ Nutzen Sie Quellen verlässlicher Informationen

Wenig Aussagekraft zur methodischen Qualität einer auf einer Internetseite angebotenen medizinischen Laieninformation haben:

- ▲ Die Angabe imposant hoch erscheinender Zugriffszahlen
- ▲ Angaben, wie viele andere Seiten auf diese Seite einen Link gelegt haben
- ▲ Benotungen der Seite durch andere Nutzer
- ▲ Verleihene Preise, so genannte Web-Awards

Weitere Informationen zur Qualität von Informationen und wie man sie erkennt, finden Sie im Internet unter der Adresse: <http://www.patienten-information.de/content/informationsqualitaet/checkliste>

Zusammengestellt von Dr. Sylvia Sänger (www.azq.de), Stand: 07/2006



Für Patienten



Toolbox
Verlässliche Informationsangebote
im Internet

Ein Service Ihres Arztes:

Liebe Patientin/Lieber Patient, wenn Sie den Wunsch haben, im Internet nach weiterführenden Informationen zu suchen, können Sie bei den nachfolgend angeführten Adressen surfen. Diese Liste soll Ihnen ein Anhaltspunkt sein und erhebt keinen Anspruch auf Vollständigkeit.

Patienteninformationsportale

- ▲ Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen bietet unter der Adresse <http://www.gesundheitsinformation.de> den Zugang zu laienverständlichen evidenzbasierten Patienteninformationen.
- ▲ Patientenfreundlich formulierte medizinische Leitlinien wie zum Beispiel die der Universität Witten-Herdecke: <http://www.patientenleitlinien.de>
- ▲ Linksammlung qualitätsgeprüfter Informationen des Ärztlichen Zentrums für Qualität in der Medizin <http://www.patienten-information.de>
- ▲ Patientenleitlinien der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften: http://www.uni-duesseldorf.de/AWMF/ll/ll_pati.htm
- ▲ Patienteninformationen der Arzneimittelkommission der deutschen Ärzteschaft: <http://www.akdae.de/45/index.html>
- ▲ Patientenleitlinien auf der Basis Nationaler Versorgungsleitlinien sind unter der Adresse: <http://www.versorgungsleitlinien.de/patienten> verfügbar.
- ▲ Die Stiftung Warentest bietet viele Informationen auch für Patienten, so zum Beispiel das Handbuch Medikamente der Stiftung Warentest (2004) <http://www.stiftung-warentest.de>
- ▲ Patienteninformationsseiten der Universität Hamburg <http://www.patienteninformation.de>
- ▲ Für Seltene Erkrankungen kann Ihr Ausgangspunkt der Suche das Angebot der Allianz Chronischer Seltener Erkrankungen sein: <http://www.achse-online.de>.

Suche nach wissenschaftlichen Veröffentlichungen/Studien

In deutscher Sprache

- ▲ Die Adresse <http://www.scholar.google.com> ist ein Serviceangebot der Suchmaschine Google. Es ermöglicht die Suche nach wissenschaftlicher Literatur.
- ▲ Das in der Schweiz ansässige Horten Zentrum für praxisorientierte Forschung und Wissenstransfer veröffentlicht unter <http://www.evimed.ch> derzeit bereits 140 Zusammenfassungen klinischer Studien zu verschiedenen Fachgebieten.

In englischer Sprache

- ▲ Themenorientierte Suchmaschine der Health on the Net-Foundation zu geprüften medizinischen Angeboten: <http://www.hon.ch/HONcode/Hunt>
- ▲ Kostenfreier Zugriff auf die Zusammenfassung von systematischen Übersichtsarbeiten der Cochrane Library: <http://www.cochrane.de/de/browse.htm>

Und wenn Sie nicht fündig wurden...

Suche mit Suchmaschinen (Bitte lesen Sie hierzu auch unser Informationsblatt zur Qualität von Patienteninformationen im Internet – Worauf Sie achten sollten!)

- ▲ Die am weitesten verbreitete Suchstrategie ist die Verwendung von Suchmaschinen. Es gibt mittlerweile hunderte von Suchmaschinen. Google <http://www.google.de> ist die bekannteste und am meisten verwendete Suchmaschine.

Zusammengestellt von Dr. Sylvia Sänger (www.azq.de), Stand: 07/2006

© ▲ Deutscher Ärzte-Verlag: R. Kunz et al., Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis, 2. Aufl.



Für Patienten

 **Toolbox**
Unterstützung des
Arzt-Patienten-Gesprächs

Ein Service Ihres Arztes:

Liebe Patientin/Lieber Patient,

Sie wurden von mir über Ihre Erkrankung und die infrage kommenden Behandlungsmethoden informiert. Sicher haben Sie hierzu noch Fragen. Bitte scheuen Sie sich nicht, diese bei Ihrem nächsten Besuch anzusprechen. Die nachfolgend aufgeführten Fragen sollen Ihnen eine Hilfe sein, noch einmal darüber nachzudenken, ob Sie alles, was ich Ihnen erzählt habe, auch wirklich verstanden haben. Natürlich können Sie auch Fragen stellen, die hier nicht angeführt sind.

- ▲ An welcher Erkrankung leide ich?
- ▲ Wie wird sich die Erkrankung auf meine persönlichen Lebensumstände auswirken?
- ▲ Welche Behandlungen kommen für meinen speziellen Fall infrage?
- ▲ Gibt es für meinen speziellen Fall verschiedene Behandlungsmöglichkeiten?
- ▲ Worin besteht der Nutzen der vorgeschlagenen Behandlung(en) für meine Lebensqualität, die Beeinflussung des Krankheitsverlaufes usw.?
- ▲ Mit welchen Risiken (Komplikationen, Nebenwirkungen) muss ich bei den vorgeschlagenen Behandlungen rechnen?
- ▲ Welche Auswirkungen hat die Behandlung auf meine persönlichen Lebensumstände und meinen Alltag?
- ▲ Mit welchen Einschränkungen in meinem persönlichen Lebensumfeld muss ich bei der Behandlung rechnen?
- ▲ Wie muss ich die verordneten Medikamente einnehmen und was ist dabei besonders zu beachten?
- ▲ Was kann ich selbst tun, um meine gesundheitliche Situation zu verbessern?
- ▲ Wo finde ich Rat und Unterstützung von Personen, die ein ähnliches gesundheitliches Problem haben wie ich?
- ▲ Wo kann ich mich über meine Erkrankung noch informieren?

Welche Fragen haben Sie außerdem noch?

Zusammengestellt von Dr. Sylvia Sängler (www.azq.de), Stand: 07/2006

©  Deutscher Ärzte-Verlag: R. Kunz et al., Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis, 2. Aufl.



Für Patienten



Toolbox
Erfahrungsaustausch mit
ebenfalls Betroffenen

Ein Service Ihres Arztes:

Liebe Patientin/lieber Patient!

Manchmal kann es hilfreich sein, sich mit anderen Personen auszutauschen, die ähnliche gesundheitliche Probleme haben. Patientenorganisationen sind hier oft ein guter Ansprechpartner, um Selbsthilfemöglichkeiten und ergänzende Unterstützungsangebote für Ihre speziellen Anliegen zu finden. Es gibt in Deutschland mehr als 70 000 Selbsthilfegruppen. Falls Sie zu Ihrer Fragestellung noch keine Patientenorganisation kennen, können Sie sich an eine der folgenden Organisationen wenden, wo man Ihnen eine kompetente Empfehlung für eine geeignete Selbsthilfeorganisation/-Gruppe geben kann:

BAG SELBSTHILFE e.V. – Bundesarbeitsgemeinschaft SELBSTHILFE von Menschen mit Behinderung und chronischer Erkrankung und ihren Angehörigen

Kirchfeldstr. 149, 40215 Düsseldorf

Telefon: 0211-31006-0, Telefax: 0211-31006-48

Email: info@bag-selbsthilfe.de, Internet: <http://www.bag-selbsthilfe.de/>

Die BAG SELBSTHILFE e.V. – Bundesarbeitsgemeinschaft SELBSTHILFE von Menschen mit Behinderung und chronischer Erkrankung und ihren Angehörigen – ist die Vereinigung der Selbsthilfeverbände behinderter und chronisch kranker Menschen und ihrer Angehörigen in Deutschland. Sie ist Dachverband von 95 (Stand 7/2006) bundesweit tätigen Selbsthilfeorganisationen, 14 Landesarbeitsgemeinschaften und drei Fachverbänden. Über ihre Mitgliedsverbände sind in der BAG SELBSTHILFE mehr als eine Million Menschen mit körperlichen, seelischen und geistigen sowie Sinnes-Behinderungen und Menschen mit unterschiedlichsten chronischen Erkrankungen zusammengeschlossen.

Allianz Chronischer Seltener Erkrankungen ACHSE e.V.

c/o DRK-Kliniken Westend, Spandauer Damm 130, 14050 Berlin

Telefon: 0180-5-ACHSE-5 oder 0180-5-22473-5

Email: info@achse-online.de, Internet: <http://www.achse-online.de>

Die ACHSE e.V. ist ein Zusammenschluss von mehr als 50 Patientenorganisationen von Kindern, Angehörigen sowie erwachsenen Betroffenen, die von einer Seltenen Erkrankung betroffen sind. Gerade für Menschen mit Seltenen Erkrankungen ist es wichtig, sich in einer Gruppe gleich Betroffener über Erfahrungen mit der Erkrankung und ihrer Behandlung auszutauschen, denn hier gibt es in den meisten Fällen noch keine Studien oder Behandlungsleitlinien. Die ACHSE setzt sich für die Verbesserung der Versorgungs- und Forschungssituation bei Seltenen Erkrankungen ein.

Der Paritätische Wohlfahrtsverband, Gesamtverband e.V.

Oranienburger Straße 13-14, 10178 Berlin

Tel: 030-24636-0, Fax: 030-24636-110

Internet: <http://www.paritaet.org/>

Das Forum chronisch Kranker und behinderter Menschen im Paritätischen Wohlfahrtsverband ist ein Zusammenschluss von Mitgliedsverbänden des Paritätischen Wohlfahrtsverbandes Gesamtverband, deren satzungsgemäße Aufgabe die Betreuung und/oder Beratung von somatisch chronisch kranken und/oder behinderten Menschen ist. Das Forum ist eine selbstorganisierte Initiative, die zugleich die Funktion eines Fachbereiches wahrnimmt. Aufgaben sind insbesondere die Förderung der fachlichen Arbeit, z.B. durch Erfahrungsaustausch, Beratung und Fortbildung, sowie die Entwicklung und Vertretung von Positionen zu gemeinsamen sozial- und gesundheitspolitischen Themen.

Sinnvoll kann es auch sein, sich an eine Patientenberatungsstelle zu wenden. Adressen sind hier zusammengestellt: <http://www.patienten-information.de/content/gesundheitsinfos/selbsthilfe>

Zusammengestellt von Dr. Sylvia Sänger (www.azq.de), Stand: 07/2006

Herausgeber- und Autorenverzeichnis

Herausgeber

Prof. Dr. med. Regina Kunz
 Institut für klinische Epidemiologie
 Universitätsspital Basel
 Hebelstrasse 10
 CH-4031 Basel

Prof. Dr. rer. nat. Dr. med.
 Günter Ollenschläger
 Ärztliches Zentrum für Qualität
 in der Medizin
 Wegelystraße 3/Herbert-Lewin-Platz
 10623 Berlin

Prof. Dr. med. Dr. phil. Heiner Raspe
 Institut für Sozialmedizin des
 Universitätsklinikums Schleswig-Holstein
 Campus Lübeck
 Beckergrube 43–47
 23552 Lübeck

Dr. med. Günther Jonitz
 Ärztekammer Berlin
 Friedrichstr. 16
 10969 Berlin

Prof. Dr. med. Norbert Donner-Banzhoff
 Philipps-Universität Marburg
 Abteilung für Allgemeinmedizin,
 Rehabilitative und Präventive Medizin
 Robert-Koch-Straße 5
 35033 Marburg

Autoren

Dr. rer. nat. Gerd Antes
 Deutsches Cochrane-Zentrum
 Institut für med. Biometrie und
 med. Informatik
 Stefan-Meier-Straße 26
 79104 Freiburg

Dr. med. Dirk Bassler, MSc (Hamilton)
 Universitätskinderklinik
 Abteilung IV (Neonatologie)
 Calwerstraße 7
 72076 Tübingen

Prof. Dr. med. Erika Baum
 Philipps-Universität Marburg
 Abteilung für Allgemeinmedizin,
 Rehabilitative und Präventive Medizin
 Robert-Koch-Straße 5
 35033 Marburg

Prof. Dr. med. Annette Becker MPH
 Philipps-Universität Marburg
 Abteilung für Allgemeinmedizin,
 Rehabilitative und Präventive Medizin
 Robert-Koch-Straße 5
 35033 Marburg

Bettina Berger
 Universität Hamburg
 MIN-Fakultät Fachwissenschaft Gesundheit
 Martin-Luther-King-Platz 6
 20146 Hamburg

Prof. Dr. med. Mathias Berger
Universitätsklinikum Freiburg
Abteilung für Psychiatrie und
Psychotherapie
Hauptstraße 5
79104 Freiburg

Martin Bergold
Evidenzbasierte Medizin Frankfurt
Institut für Allgemeinmedizin
Universitätsklinikum Frankfurt/Main
Theodor-Stern-Kai 7
60590 Frankfurt/Main

Dr. med. Michael M. Berner
Universitätsklinikum Freiburg
Abteilung für Psychiatrie und
Psychotherapie
Hauptstraße 5
79104 Freiburg

Dr. rer. nat. Hilke Bertelsmann
Gemeinsamer Bundesausschuss
Stabsstelle Methodik und Information
Auf dem Seidenberg 3a
53721 Siegburg

Prof. Dr. med. Dr. phil.
Nikola Biller-Andorno
Lehrstuhl für Biomedizinische Ethik
Ethik-Zentrum Universität Zürich
Zollikerstrasse 115
CH-8008 Zürich

Dr. med. Wolfgang A. Blank
Facharzt für Allgemeinmedizin/
Gemeinschaftspraxis
Ferdinand Neumaier Straße 6
94259 Kirchberg im Wald

Dipl. Biol. Insa Bruns
Geschäftsstelle des Netzwerks der
Koordinierungszentren für Klinische
Studien
c/o IMSIE-Gebäude 22
Kerpener Straße 62
50937 Köln

Dr. med. Johannes Bruns
Deutsche Krebsgesellschaft e.V.
Geschäftsstelle
Steinlestraße 6
60596 Frankfurt am Main

Dr. rer. medic Frank Brunsmann
BMG-Projekt „Patientenselbsthilfe und
Versorgungsqualität bei seltenen
Erkrankungen“
Charité-Augenklinik
Campus Virchow-Klinikum
Augustenburger Platz 1
13353 Berlin

Prof. Dr. med. Heiner C. Bucher
Institut für klinische Epidemiologie
Universität Basel
Hebelstrasse 10
CH-4031 Basel

Dr. med. Gerd Burmester
Institut für Klinische Pharmakologie
Klinikum Bremen-Mitte gGmbH
Sankt-Jürgen-Straße 1
28177 Bremen

Prof. Dr. med. Reinhard Busse MPH, FFPH
Technische Universität Berlin
Management im Gesundheitswesen
Straße des 17. Juni 145
10623 Berlin

Prof. Dr. rer. biol. hum. Marie-Luise Dierks
Medizinische Hochschule Hannover
Abteilung für Epidemiologie, Sozialmedizin
und Gesundheitssystemforschung
OE 5410
30623 Hannover

Prof. Dr. med. Norbert Donner-Banzhoff
Philipps-Universität Marburg
Abteilung für Allgemeinmedizin,
Rehabilitative und Präventive Medizin
Robert-Koch-Straße 5
35033 Marburg

Günther Egidi
Arzt für
Allgemeinmedizin/Gemeinschaftspraxis
Huchtinger Heerstraße 41
28259 Bremen

Prof. Dr. med. Yngve Falck-Ytter
Case Western Reserve University
Louis Stokes VA Medical Center Cleveland
Division of Gastroenterology
10701 East Blvd
Cleveland, Ohio 44106, USA

Prof. Dr. med. Johannes Forster
St. Josef-Krankenhaus
Pädiatrie
Abteilung für Kinder- und Jugendmedizin
St. Hedwig
Sautier Straße 1
79104 Freiburg

Dr. med. Bernhard Gibis, MPH
Kassenärztliche Bundesvereinigung
Dezernat 2
Versorgungsqualität und Sicherstellung
Herbert-Lewin-Platz 2
10623 Berlin

Prof. Dr. med. Johannes Giehl
Kompetenz-Centrum
Qualitätssicherung/Qualitätsmanagement
KCQ
MDK Baden-Württemberg
Konrad-Adenauer-Straße 15
72072 Tübingen

Dr. rer. nat. Britta Göhlen, PhD
DIMDI-Deutsches Institut für Medizinische
Dokumentation und Information
Deutsche Agentur für Health Technology
Assessment/DAHTA
Waisenhausgasse 36–38a
50676 Köln

Prof. Gordon Guyatt MD, MSc
Clinical Epidemiology and Biostatistics and
Medicine
Faculty of Health Sciences
McMaster University
Health Sciences Centre, Room 2C12
1200 Main Street West
Hamilton, Ontario, L8N 3Z5
Canada

Prof. Dr. jur. Dieter Hart
Institut für Gesundheits- und Medizinrecht
Fachbereich Rechtswissenschaft
Universität Bremen
Postfach 33 04 40
28334 Bremen

Prof. Dr. med. Dr. phil. Martin Härter
Universitätsklinikum Freiburg
Abteilung für Psychiatrie und
Psychotherapie
Sektion Klinische Epidemiologie und
Versorgungsforschung
Hauptstraße 5
79104 Freiburg

Dr. med. Hanna Kaduszkiewicz
Universitätsklinikum Hamburg-Eppendorf
Martinistraße 52
20246 Hamburg

Heidmarie Keller
Philipps-Universität Marburg
Abteilung für Allgemeinmedizin,
Rehabilitative und Präventive Medizin
Robert-Koch-Straße 5
35033 Marburg

Dr. med. Regina Klakow-Franck
Bundesärztekammer
Dezernat 3: Qualitätssicherung
Herbert-Lewin-Platz 1
10623 Berlin

Prof. Dr. med. Johannes Köbberling
Zentrum für Innere Medizin
Kliniken St. Antonius Wuppertal
Carnaper Straße 48
42283 Wuppertal

Dr. rer. biol. hum. Inke R. König
Institut für Medizinische Biometrie und
Statistik
Universität zu Lübeck
Ratzeburger Allee 160, Haus 4
23538 Lübeck

PD Dr. med. Ina Kopp
Leitlinien AWMF
Klinikum der Philipps-Universität Marburg
Zentrum für Operative Medizin
Institut für Theoretische Chirurgie
Baldinger Straße 1
35043 Marburg

Dr. med. Christoph Kreck
Sozialmedizinische Expertengruppe
Methoden- und Produktgruppe (SEG7) der
MDK-Gemeinschaft beim Medizinischen
Dienst der Spitzenverbände der
Krankenkassen
Lützowstraße 53
45141 Essen

Dr. med. Tanja Krones
Dipl. Soziologin
Bioethik – Klinische Evidenz
Fachbereich Medizin
Klinikum der Philipps-Universität Marburg
Baldingerstraße
35032 Marburg

Dr. med. Thomas Kühlein
Universitätsklinikum Heidelberg
Abteilung Allgemeinmedizin und
Versorgungsforschung
Voßstraße 2
69115 Heidelberg

Prof. Dr. med. Regina Kunz
Institut für klinische Epidemiologie
Universitätsspital Basel
Hebelstrasse 10
CH-4031 Basel

Prof. Dr. rer. nat. Walter Lehmacher
Institut für medizinische Statistik,
Informatik und Epidemiologie
Universitätsklinikum Köln
Kerpener Straße 62
50937 Köln

Prof. Dr. oec. publ. Reiner Leidl
Institut für Gesundheitsökonomie und
Management im Gesundheitswesen
GSF-Forschungszentrum für Umwelt und
Gesundheit
Postfach 1129
85758 Neuherberg

Monika Lelgemann, MSc
HTA Zentrum in der Universität Bremen
c/o Institut für Gesundheits- und
Medizinrecht
Universität Bremen
Postfach 330440
28334 Bremen

Dr. phil. Christian Lenk
Abteilung für Ethik und Geschichte der
Medizin
Universität Göttingen
Humboldtallee 36
37073 Göttingen

Dr. med. Guido Lerzynski, MSc
Diözesan-Caritasverband Köln
Bereich Krankenhäuser und Altenhilfe
Georgstraße 7
50676 Köln

Prof. Dr. med. Markus Löffler
Kordinierungszentrum
für Klinische Studien Leipzig
Universität Leipzig
Härtelstraße 14–16
04107 Leipzig

Dr. med. Dagmar Lühmann
Medizinische Universität Lübeck
Institut für Sozialmedizin
Beckergrube 43–47
23552 Lübeck

Prof. Dr. med. Herbert Maier-Lenz
Zentrum für Klinische Studien Freiburg
Universitätsklinikum Freiburg
Elsässer Straße 2
79110 Freiburg

Dr. phil. Thorsten Meyer
Institut für Sozialmedizin des
Universitätsklinikums Schleswig-Holstein
Campus Lübeck
Beckergrube 43–47
23552 Lübeck

Dr. med. Karl-Heinz Moser
Chirurgische Gemeinschaftspraxis Südstadt
Karolingerring 31
50678 Köln

Univ.-Prof. Dr. med. Ingrid Mühlhauser
Universität Hamburg
MIN-Fakultät Fachwissenschaft Gesundheit
Martin-Luther-King-Platz 6
20146 Hamburg

Prof. Dr. rer. nat. Dipl.-Ing.
Edmund A. M. Neugebauer
IFOM-Institut für Forschung in der
operativen Medizin
Lehrstuhl für Chirurgische Forschung
Private Universität Witten/Herdecke
gGmbH
Ostmerheimer Straße 200
51109 Köln

Prof. Dr. rer. nat. Dr. med.
Günter Ollenschläger
Ärztliches Zentrum für Qualität
in der Medizin
Wegelystraße 3/Herbert-Lewin-Platz
10623 Berlin

PD Dr. med. Matthias Perleth, MPH
AOK-Bundesverband
Stabsbereich Medizin
Dependance Berlin
Charlottenstraße 42
10117 Berlin

Dr. med. Uwe Popert
Arzt für Allgemeinmedizin
Dörnbergstraße 21
34119 Kassel

Dr. rer. nat. Bernd Quadder
Deutsche Sarkoidose Vereinigung e.V.
Bundesverband
Postfach 3043
40650 Meerbusch

Dr. med. Josef Ramsbacher
Arzt für Neurochirurgie
Schlüterstraße 38
10629 Berlin

Prof. Dr. med. Dr. phil. Heiner Raspe
 Institut für Sozialmedizin des
 Universitätsklinikums Schleswig-Holstein
 Campus Lübeck
 Beckergrube 43–47
 23552 Lübeck

Dr. med. Martin Ruß
 Institut für Medizinische Immunologie
 Klinikum der Medizinischen Fakultät
 Martin-Luther-Universität Hall-Wittenberg
 Magdeburger Straße 2
 06097 Halle (Saale)

Dr. med. Alric Rütther
 DIMDI-Deutsches Institut für Medizinische
 Dokumentation und Information
 Deutsche Agentur für Health Technology
 Assessment/DAHTA
 Waisenhausgasse 36–38a
 50676 Köln

Eva-Maria Sadowski
 Philipps-Universität Marburg
 Abteilung für Allgemeinmedizin,
 Rehabilitative und Präventive Medizin
 Robert-Koch-Straße 5
 35033 Marburg

Dr. PH Sylvia Sängler
 Bereichs Medizinische Information
 Ärztliches Zentrum für Qualität in der
 Medizin
 Wegelystraße 3/Herbert-Lewin-Platz
 10623 Berlin

Dr. med. Stefan Sauerland
 IFOM-Institut für Forschung in der
 operativen Medizin
 Lehrstuhl für Chirurgische Forschung
 Private Universität Witten/Herdecke
 gGmbH
 Ostmerheimer Straße 200
 51109 Köln

Prof. Dr. med. Torsten Schäfer, MPH
 Institut für Sozialmedizin des
 Universitätsklinikums Schleswig-Holstein
 Campus Lübeck
 Beckergrube 43–47
 23552 Lübeck

Dr. med. Nicole Schlottmann
 Dezernat V Medizin
 Deutsche Krankenhausgesellschaft
 Wegelystraße 3
 10623 Berlin

Prof. Dr. med. Matthias Schneider
 Rheumazentrum Düsseldorf
 Postfach 10 10 07
 40001 Düsseldorf

Prof. Dr. Dr. med. Holger Schünemann Md,
 MSc, PhD
 Clinical Research Development and
 Information Translation (INFORMA) Unit
 Italian National Cancer Institute Regina
 Elena (IRE)
 Via Elio Chianesi 53
 I-00144 Rom

Prof. Dr. med. Günter Seidel
 Klinik für Neurologie des
 Universitätsklinikums Schleswig-Holstein
 Campus Lübeck
 Ratzeburger Allee 160
 23538 Lübeck

Dr. med. Christoph M. Seiler, MSc
 Studienzentrum der Deutschen Gesellschaft
 für Chirurgie
 Universitätsklinikum Heidelberg
 Im Neuenheimer Feld 110
 69120 Heidelberg

Dr. med. Nicole Skoetz
 Koordinierungszentrum für klinische
 Studien
 Gleueler Straße 88a
 50931 Köln

Prof. Dr. med. Andreas C. Sönnichsen
Institut für Allgemein-, Familien- und
Präventivmedizin
Strubergasse 21
A-5020 Salzburg

Dr. med. Frank Thalau MPH, MSc
Charité – Universitätsmedizin Berlin
Institut für Sozialmedizin, Epidemiologie
und Gesundheitsökonomie
Luisenstraße 57
10117 Berlin

Henning Thole
Verband Forschender
Arzneimittelhersteller e.V.
Gesundheitsökonomie –
Leitlinienentwicklung
Hausvogteiplatz 13
10117 Berlin

Dr. med. Christian Thomeczek
Ärztliches Zentrum für Qualität in der
Medizin
Wegelystraße 3/Herbert-Lewin-Platz
10623 Berlin

PD Dr. med. Antje Timmer, MSc
Deutsches Cochrane-Zentrum
Abteilung Med. Biometrie und Statistik
Klinikum der Albert-Ludwigs-Universität
Stefan-Meier Straße 26
79104 Freiburg

Prof. Dr. med. dent. Jens C. Türp
Klinik für Rekonstruktive Zahnmedizin und
Myoarthropathien
Zentrum für Zahmedizin der Universität
Basel
Hebelstrasse 3
CH-4056 Basel

Dr. Claus Wächtler
Abteilung für Psychiatrie und
Psychotherapie – Gerontopsychiatrie
Asklepios Klinik Nord – Ochsenzoll
Langenhorner Chaussee 560
22419 Hamburg

Dr. med. Robert von Wattenwyl
Unidad Cirurgia Cardiovascular de
Guatemala (UNICAR)
9 Avenida
8-00 Zona 11
01002 Guatemala City, Guatemala

Dr. med. Tobias Weberschock
Evidenzbasierte Medizin Frankfurt
Institut für Allgemeinmedizin
Universitätsklinikum Frankfurt/Main
Theodor-Stern-Kai 7
60590 Frankfurt am Main

Prof. Dr. rer. pol. Karl Wegscheider
Wegscheider Biometrie und Statistik GmbH
Barstraße 9
10713 Berlin

Dr. med. Susanne Weinbrenner
Ärztliches Zentrum für Qualität in der
Medizin
Wegelystraße 3/Herbert-Lewin-Platz
10623 Berlin

Dr. med. Olaf Weingart
Cochrane Haematological Malignancies
Group
Klinik I für Innere Medizin der Universität
zu Köln
Kerpener Straße 62
50924 Köln

PD Dr. jur. Felix Welti
Konrad-Adenauer-Straße 3
23558 Lübeck

Prof. Dr. med. Karl Werdan
Klinik und Poliklinik für Innere Medizin III
Martin-Luther-Universität
Ernst-Grube-Straße 40
06120 Halle/Saale

Dr. med. Hans Wille
Institut für Klinische Pharmakologie
Klinikum Bremen-Mitte gGmbH
Sankt-Jürgen-Straße 1
28177 Bremen

Prof. Dr. med. Jürgen Windeler
Medizinischer Dienst der Spitzenverbände
der Krankenkassen e.V.
Lützowstraße 53
45141 Essen

Stephanie Wolff, M.A.
Geschäftsstelle des Netzwerks der
Koordinierungszentren für Klinische
Studien
c/o IMSIE-Gebäude 22
Kerpener Straße 62
50937 Köln

Prof. Dr. med. Hans-Reinhard Zerkowski
Ordinarius für Herz- u. Thoraxchirurgie
Universität Basel
Auf der Bischoffhöhe 6
4125 Riehen, Schweiz

Prof. Dr. rer. nat. Andreas Ziegler
Institut für Medizinische Biometrie und
Statistik
Universität zu Lübeck
Ratzeburger Allee 160, Haus 4
23538 Lübeck

Stichwortverzeichnis

2-Test 145

A

Abstract 253
 Abwägung 406
 Abweichung, systematische 487
 Acetyl-Cholinesterasehemmer 271
 Adhärenz 33
 AGREE 178, 493
 AGREE Collaboration 184
 Akkreditierungen 426
 Akzeptanz 47
 Algorithmus 365, 368
 Allgemeinmedizin 289
 Alltagsproblem 47
 Alter 405
 Analogskala, visuelle 209
 Angemessenheit, inhaltliche 183
 Angemessenheits-Forschung 64
 Anwendbarkeit 177
 Approbationsordnung 421
 Appropriateness Research 64
 Arbeitszufriedenheit 31
 ARRIBA-Herz 295
 Arthritis, frühe rheumatoide 284
 Arzneimittel 479
 Arzneimittelprüfungen 469
 Arzthaftungsprozess 393
 Arzthaftungsrecht 393
 Ärztliches Zentrum für Qualität in der Medizin 457
 Arzt-Patienten-Kommunikation 217
 Arzt-Patienten-Verhältnis 52
 ASSERT 493
 Assoziation 115f.
 Attribution-Bias 334
 Aufklärungspflichtverletzung 393
 Aus-, Fort- und Weiterbildung 424, 441
 Außenbandrupturen 267
 Auswahl-Bias 57

B

Bandscheibenvorfall 319
 – lumbaler 319

Basisrisiko 78
 Bedeutung, klinische 90
 Bedside-Teaching 414
 Behandlungsfehler 393, 397
 Behandlungsleitlinien 336
 Behandlungsnutzen
 – Überschätzung 153
 Behandlungspfade 286
 Behandlungsprogramme bei chronischen
 Krankheiten, strukturierte 404
 Beobachtungsstudien 62, 237, 239
 Beratung, ärztliche 37
 Beratungssituation, hausärztliche 300
 Beratungsstrategie 295
 Berechnungsinstrument 302
 Berichtsqualität 182
 Beurteilungskriterien 178
 – für Leitlinien 462
 Bevölkerungs-Efficacy 66
 Beweiskraft 40
 Bewertung, kritische 42, 90, 177, 340
 Bewertungsinstrumente 177
 Bias 390, 415, 484, 486f.
 – Selection 151
 Bildungsniveau 97
 Bioresonanztherapie 408
 Brustkrebs 101, 278
 Brustkrebsrisiko, individuelles 278
 Bundesärztekammer 457
 Bundesministerium für Gesundheit 376
 Bundessozialgericht 401, 407
 Bundesverfassungsgericht 401

C

Case-Management 338
 CASP 429, 493
 Cauda-equina-Syndrom 321
 Chalmers 443
 Checkliste 177, 336, 362, 491
 Chemonukleolyse 321
 Chlorpropamid-Alkohol-Flush-Test 7
 Cholinesterasehemmer 273
 Clearing-Stellen 182
 Clearing-Verfahren 182

Clinical Evidence 251, 253
 Cochrane Collaboration 443
 Cochrane Library 219, 339, 441, 443, 476
 Cochrane, Archibald 443
 Compliance 79
 Concealment 485
 Confounder 79, 103, 483f., 486
 Confounding 79, 240, 484
 – durch Ethnizität 117
 CONSORT 492
 Consumer 451
 Critical Appraisal 42, 183
 Critical-appraisal-skills-Programm 429
 Curriculum
 – Evidenzbasierte Medizin 424, 440
 – Patientenberatung 431

D

Darmkrebs 278
 Datenbanken 475
 – medizinische 475
 Datenbankrecherche 476
 Defizite, motorische 325
 DELBI 177, 362, 493
 Delphi-Methode 64
 Demokratie 405
 Denken, ärztliches 261
 Denominator 97
 Depression 336
 Depressive Erkrankungen 335
 Detection signal bias 98
 Detektionsbias 111
 Deutsche Zentralbibliothek für Medizin 477
 Deutsches Cochrane Zentrum 443
 Deutsches Institut für Medizinische Dokumenta-
 tion und Information 475
 Deutsches Leitlinienbewertungs-Instrument
 362
 Deutsches Leitlinien-Clearingverfahren 459
 Diagnose 262
 Diagnostischer Test
 – carcinoembryonales Antigen (CEA) 123
 – Goldstandard 122
 – klinische Bedeutung 129
 – Patientenspektrum 123
 – Reproduzierbarkeit 124, 129
 – Validität 122
 Die klassischen Fünf Schritte 263
 Dienste, medizinische 345
 Differential Missings 79
 Differenzialdiagnose 94f.
 DIMDI 447, 475

Dissektomie 321
 Disease orientated evidence (DOE) 41, 250
 Diskontierung 208
 Diskursanalyse 167
 Diskussion, kollegiale 264
 DNEbM 439
 Dokumentation, juristisch-gutachterliche 270
 Dokumentationspflicht 396
 Domne 178
 Donepezil 271
 Dosis-Wirkungsbeziehung 111, 240
 Duchenne'sche Muskeldystrophie 408

E

EbM im Studium 419, 440
 EbM und Medical Professionalism 22
 EbM-Kurskonzept, klassisches 422
 EbM-Laienschulung 429
 EbM-Rezeption 264
 EbM-Strategie 364
 Educational Prescription 414
 Effekt, präventiver 299
 Effektgröße 45, 62
 Effektivität 62, 341
 Effektmaß 489
 Effektstärken 367
 Effizienz 203, 390
 Einbeziehung des Patienten 300
 Einflussgröße 79
 Einschlusskriterien 96
 Einzelfall 82
 Einzelfallstudien 81
 Elektrostimulation, funktionelle 327
 EMBASE 476
 Empfehlung 180, 231, 244
 – methodische 345
 – versorgungsübergreifende 286
 Empfehlungsgrad 180, 365
 Empfehlungsstärke 236f.
 Endpunkt 236, 238, 243
 – harter 37
 – patientenrelevanter 52
 – primärer 34, 141
 Entscheiden 261
 Entscheidung, informierte 278
 Entscheidungen 61, 76
 Entscheidungsfindung 361
 – partizipative 340
 Entscheidungsgrundlage 52
 Entscheidungshilfen 361, 371
 Entscheidungskontext 211
 Entscheidungsprozess 39

Entscheidungsregel 262
 – klinische 269
 Entscheidungszwang 84
 Entwicklung, curriculare 431
 Epidemiologie
 – genetische 115
 – klinische 491
 Erfahrung, klinische 273
 Erfassungsqualität 108
 Erfolgserlebnisse 31
 Erkenntnis 3f., 45, 77
 Erkenntnisgewinn 4
 Erkrankungen, metabolische 446
 Erlaubnisvorbehalt 385
 Ethik der Evidenzbasierten Medizin 24
 EUebm-Projekt 425
 EUNetHTA 201
 European Network for Health Technology
 Assessment 201
 Evaluation 341
 – formative 418
 – ökonomische 205
 – summative 418
 Evidence based health care 61, 378
 Evidence based medicine tool kit 493
 Evidenz 51, 75, 218
 – Abgleich 39
 – best verfügbare 181
 – externe 39, 61, 293
 – methodisch aufgearbeitete 367
 – ökonomische 211f.
 – unmittelbare 332
 – Verwenden 39
 Evidenzberichte 252, 371, 457
 Evidenzeinstufung 180
 Evidenzfindung 75
 Evidenzgenerierung 75
 Evidenzgrad 40, 365
 Evidenzhierarchie 79
 Evidenzklassen 57
 Evidenzklassensysteme 76
 Evidenzprofil 241
 Evidenz-Shifting 180
 Evidenzstärke 365
 Evidenzstufen 135
 Evidenztabellen 365
 Evidenzvermittlung 52
 Experimental Event Rate = EER 141
 Expertengruppen 183
 Expertenkreise 458
 Expertise 39
 – klinische 39

F

Fachverbände 405
 Falldefinition 108
 Fall-Kontroll-Studie 101
 Fallserien 136
 Fallzahlberechnung 99
 Fallzahlplanung 490
 Falsifikation 4
 Familie 331
 Fehler, systematische (Bias) 135
 Forcierter Gebrauch (Constraint Induced (CI)
 Movement Therapy, sog. Taub-Training)
 326
 Förderung über § 65b SGB V 431
 Formeln, multivariate 296
 Forschung 44, 467
 – klinische 16
 Forschungsbedarf 468
 Forschungsparadigma 159
 Fortbildung, ärztliche 424
 Frage
 – patientenzentrierte 89
 – relevante 89
 – suchtaugliche 242
 Fragestellung 89, 364
 – ätiologische 101
 – fokussierte 241
 – konkret 41
 – präzise medizinische 414
 Framing 56
 Framingham-Formel 296

G

Galantamin 271
 Gemeinsamer Bundesausschuss 199, 212, 376,
 403
 Generalisierbarkeit 79
 Generalist 42
 Genotypisierungsfehler 118
 Gesamttafel 78
 Geschichte der EbM 20
 Geschlecht 405
 Gesprächsführung 300
 Gesundheitssystem 61, 76
 Gesundheitsversorgung, Evidenzbasierte 61
 G-I-N 184
 – Guideline Library 185
 GKV 380
 Goldstandard 77
 – diagnostischer Test 122
 GRADE Working Group 231
 GRADE-Methodik 232

GRADE-Profiler 241
 Graduierung 180
 Grenzen Evidenzbasierter Medizin 289
 Grundkurs 425
 Grundpopulation 135
 Grundsatzgutachten 346
 Guideline Clearinghouses 182
 Guidelines International Network 184

H

Haftungsrecht 393
 Hämatologie 446
 Handlungsalgorithmus, fünfstufiger 336
 Handlungsempfehlungen 178, 232, 361, 371
 Haplotyp 118
 Hardy-Weinberg-Gleichgewicht 118
 Hausärzte 42
 Health Technology Assessment 191, 475
 Heilversuche 408
 Herzerkrankung, koronare 305
 Herz-Kreislauf-Prävention 295
 Heterogenität
 – Meta-Analyse 154
 Heterogenität der Ergebnisse 238
 Heuristik 262
 Hidden Bias 79
 Hierarchie des Wissens 387
 Hormonersatz-Therapie 101
 Horten-Zentrum 252
 Hospitalisierungsrate 333
 HTA 191, 475
 – international 199
 HTA-Berichte 71, 477
 HTAi 199
 Huckepackeffekte 79
 Hypericum-Extrakt (Johanniskraut) 335

I

Impact von EbM 264
 Implementierung 283, 369
 Implementierungshindernis 181, 283
 Implementierungsmaßnahmen 287
 INAHTA 200
 Indikation 46
 Indikationsstellung 112
 Indikatoren 361
 Induktion 164
 Information 395, 475
 – Evidenzbasierte 217, 278
 – externe 263
 Informations- und Kommunikationsplattform 439

Informationsanbieter 475
 Informationsangebot 256
 Informationsbedarf 256
 Informationsmanagement 250
 Informationssystem 255, 475
 Informationszeitalter 264
 Informed Consent 277
 Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen 404
 Integrated Teaching 425
 Intention to treat 272, 485
 Intention to treat-Analyse 139
 Interaktionseffekte 79
 Interaktionstest 79
 International Guideline Library 185
 International Network for Agencies for HTA 200
 Internet 84
 Internetrecherche 415
 Interventionsstrategien 337
 Interviewer-Bias 109
 Inzidenz 104
 IQWiG 398

J

Journal Club 252, 418, 423

K

kappa-Statistik 153
 Kassenärztliche Bundesvereinigung 457
 Kassenärztliche Vereinigungen 403
 Kausalität 111
 Kinder 331
 KKS 467
 Klassifikationen 480
 Klassifizieren 232, 238
 Kleingruppen-Arbeit 422
 Knock-Out-Kriterien 181
 Kohortenstudie 101, 136
 Kombra-Projekt 434
 Kommunikation der Evidenz 55
 Kompetenztraining für Patienten- und VerbrauchervertreterInnen 432
 Konfidenzintervall 99, 104, 144, 485, 487, 489
 Konfidenzwahrscheinlichkeit 487
 Konsens 12
 Konsensfindung 362, 368
 Konstruktivismus 163
 Kontinuum 261
 – kognitives 261
 Kontrolle 81
 Kontrollgruppe 83, 103

Koordinierungszentren für Klinische Studien 467
 Kopplungs-Ungleichgewicht 116
 Kosten 206
 – direkte 207
 – indirekte 207
 Kosteneffektivität 68, 70
 Kosteneffektivitäts-Analyse 206
 Kosteneffektivitäts-Relation 210
 Kosten-Minimierungs-Analyse 206
 Kosten-Nutzen-Analyse 206
 Kostenübernahme-Entscheidung 192
 Krankenhäuser 403
 Krankenversicherung 401
 Krankheit, iatrogene 8
 Krebsfrüherkennungs-Programm 280
 Kreuzschmerzen 93
 Kurse für Fortgeschrittene 425
 Kyphoplastie 345

L

Laien 427, 451
 Last observation carried forward 272
 LEAD 428
 Leading Education and Advocacy Development 428
 Lebensqualität 52, 405
 Lebensqualität, gesundheitsbezogene 208
 Lehr- und Lernzielkatalog 424
 Lehr- und Moderatorentaining 420
 Lehre, Evidenzbasierte 413
 Lehrerfolg 419
 Leistungserbringungsrecht 403
 Leistungskatalog 213
 Leitlinien 32, 48, 177, 231, 290, 361, 393, 457
 – Evidenzbasierte 404
 – Methodik 362
 – Umsetzung 362
 – Ziele 361
 Leitlinienanwender 362
 Leitlinienbewertung 177
 Leitliniendiseminierung 370
 Leitlinien-Empfehlung des Europarates 184
 Leitlinienentwicklung 335, 363, 368
 Leitliniengruppe 365
 Leitlinienmanual von AWMF und ÄZQ 184
 Leitlinienqualität 182
 Leitlinien-Qualitätsprogramm 336
 Leitlinienreport 369
 Leitlinisierung 389
 Leonardo da Vinci 425
 Lernen, problemorientiertes 422

Lernerfolg 419
 Lerntools 425
 Lernzieldefinitionen 425
 Level of Evidence 365
 Likelihood-Ratio 269
 Literatur 480
 – Auswahl 365
 – wissenschaftliche 475
 Literaturlaufarbeitung 365
 Literaturrecherche 338, 365
 Literatursuche 90
 Living Guideline 181
 Lokomotionstherapie 327
 Lost to follow-up 138
 Lösungsstrategien 89
 Lumbalgie 289, 321

M

Mammographie-Screening 61, 277
 Mängel, methodische 275
 Matching 80, 103
 MDK 347
 Medikamenteneffekt 78
 Medizinethik 387
 Medizininformation 475
 Medizinprodukte 475
 Medizinrecht 394
 Medizinstudium 419
 Medizinwissen 475
 MEDLINE 448, 476
 Meinung, konsentiertere 219
 Meinungsbildner 426
 Merkblatt 332
 Messgrößen 370
 Meta-Analyse 51, 75, 219, 232, 339, 445
 – fokussierte Fragestellung 151
 – Grenzen 149
 – Klinischer Nutzen 155
 – statistische Zusammenfassung 153
 Methode der zeitlichen Abwägung 209
 Methodenlehre 76
 Methodenreport 182
 Methodik 45
 Methodologie 165
 Mindlines 254
 Mittelwerte 487
 Modellierung 204
 Monitoring 473
 Morbidität 405
 Mortalität 405
 Multidisziplinarität 181, 183
 Multiplikatoren-Programm 420

- Myoarthropathien 351
 - Autoritäten 354
 - charismabasierte Behandlungsphilosophien 356
 - Evidenzbasierte Zahnmedizin 357
 - Praxisvariation 351
 - Über-, Unter- und Fehlversorgung 351

N

- Nachbeobachtung 98
- Nachtest-Wahrscheinlichkeit 124
- National Institute for Health and Clinical Excellence [NICE] 212
- Nationale VersorgungsLeitlinie 365, 459
- Neurochirurgie 322
- NICE 389
- Nichtteilnehmerquote 108
- Niedergelassener 31
- Niedrigprävalenz 292
- Niedrig-Prävalenz-Bereich 35
- NNT 487, 490
- Norm 394
- Notfallambulanzen 332
- Nullhypothese 489
- Number needed to harm (NNH) 36
- Number needed to treat (NNT) 36, 264
- Nutzen
 - diagnostischer und therapeutischer 403
 - erwarteter 55
- Nutzen-Risiko-Abwägung 112, 367
- NVL-Programm 184

O

- Odds-Ratio 118, 153, 487
- Öffentlichkeit 405
- Onkologie 446
- Organisationspflicht 396
- Organisationspflichtverletzung 394
- Osteoporose 345
- Ottawa Ankle Rule 267
- Outcome 79, 232, 240
- Overt Bias 79

P

- Paper-Cases 414
- Parameter der Relevanz 45
- Patient oriented evidence that matters 41
- Patienten 427
 - depressiv 335
- Patientenauswahl 95
- Patientenautonomie 387
- Patientenberater 431

- Patientenbroschüren 278
- Patienten-Compliance 67
- Patientenebene 275
- Patientengruppen 389
- Patienteninformation 277, 457
 - Evidenzbasierte 218
- Patientenorganisationen 58
- Patientenorientierte RECHERCHE 423
- Patientenperspektive 51, 233
- Patientenpräferenzen 234
- Patientenvertreter 428
- Patientenworkshops 431
- Peer-Review 152
- Per protocol- oder "as treated"-Analyse 139
- PERLEN 423
- Pflegeversicherung 403
- Placeboeffekt 137
- Plattform
 - Information 439
- Plausibilität, biologische 155
- Pluralismus 77
- POEM (Patient orientated evidence that matters) 250
- Politikfeldanalyse 191
- Popper, Karl 4
- Populations-Stratifikation 117
- Power 79, 490
- Präferenzen des Patienten 52
- Prätest-Wahrscheinlichkeit 93
- Prävalenz 93, 95, 126
- Prävention 16, 277
- Präzision 239, 243
- Precision of the estimate 140
- Pregnancy and Childbirth 445
- Primär-Versorgung 332
- Probleme, identifizierte 91
- Problemlösung 89
- Problemlösungs-Pyramide 263
- Professionalisierung 287
- Prognose 95, 98f.
- Programm für Nationale VersorgungsLeitlinien 184
- Prozess- und Ergebnisqualität 362
- Prüfung 418
- Pseudokrapp 331
- Psychiatrie 335
- Psychosomatik 335
- Psychotherapie 335
- Publication-Bias 150, 152
- PubMed-Recherche 34
- p-Wert 144

Q

QUADAS 492
 Qualität 361, 403
 – der Evidenz 232, 234, 238ff.
 Qualitätsbewertung 231
 Qualitätsentwicklung in der Medizin 457
 Qualitätsindikatoren 361, 370
 Qualitätskriterium 178
 Qualitätsmanagement 361, 396, 421
 Qualitätsmerkmal 178
 Qualitätssicherung 396, 457
 Qualitätsstandard 467
 Qualitätszirkel 33, 177, 332, 423
 Quality-adjusted life years, QALYs 209
 Quellen 75
 QUORUM 492

R

Randomisation 80
 Randomisierter, kontrollierter Versuch (RCT) 109
 Randomisierung 137, 485
 RCT 62, 81, 137, 220, 365
 Recall-Bias 103, 136
 Recherche 33, 365, 476
 Rechtsprechung 398
 Rechtsstaat 405, 407
 Red Flags 290
 Reduktion der Röntgendiagnostik 269
 Referenztest, s. *Goldstandard*
 Reflection in action 263
 Reflektieren, ärztliches 264
 Reflexion, kritische 263
 Register 80
 Regressionsanalysen, logistische 109
 Rehabilitation 403
 – robot-assistierte 326
 Rekrutierung 95
 Relevanz 41, 177, 340, 416, 490
 – klinische 243
 Replicating Mode 263
 Reporting-Bias 239
 Responder 273
 Ressourcen, finanzielle 68
 Ressourcenallokation 387
 Restriktion 109
 Revascularisation 305
 – interventionell 307, 309
 – operativ 307
 Revascularisationsmethode 307
 Review, systematischer 62, 76, 133, 232, 468

Review-Gruppen 339
 Richtlinien
 – Bundesausschuss 403
 Risiko
 – absolutes 35
 – relatives 35, 104, 141, 487
 Risikofaktoren 137
 – kardiovaskuläre 295
 Risikomanagement 396
 Risikomanagementsystem 393
 Risikoprognose 295
 Risikorechner 278
 Risiko-Reduktion 34, 45
 – absolute 35
 – relative 35, 264
 Risikoschätzer
 – absolute 141
 – relative 141
 Risk-Ratio 153
 Rivastigmin 271
 Routine 262

S

S3-Leitlinie 285
 Sachverständige 393, 407
 Schätzen 485
 Schlaganfall 325
 Schnittstellen 394
 Schnittstellenprobleme 286
 Schulen, allgemeinbildende 434
 Schulmedizin 409
 Schulung 428
 Schulungsprogramme für Laien 434
 Schutzpflicht des Staates für Leben und Gesundheit 408
 Screening 110, 277
 Screening-Programm 61
 Screening-Test
 – carcinoembryonales Antigen (CEA) 123
 Searching Mode 263
 Sekundärprävention 299
 Selbstbestimmungsaufklärung 396
 Selbsthilfegruppen 431
 Selbststudium 424
 Selbsttäuschung 5, 77
 Selbstverwaltungskörperschaft 457
 Selbstverwaltungsorganisation 457
 Selection-Bias 103, 136
 Seminar 414
 Sensitivität 124, 487
 – von etwa 98% 269

Sensitivität/Spezifität
 – Testausbeute 129
 – und Wahrscheinlichkeitsverhältnis 130
 Sensitivitätsanalyse 211
 – Meta-Analyse 154
 Shared decision making 217, 340
 Sicherheitsdenken 270
 Sicherheitsprobleme 80
 Sicherungs- oder therapeutische Aufklärung
 395
 Signifikanztest 485
 Simpsons Paradox 78
 SnNout 128
 Sozialmedizin 345
 Sozialrecht 401
 Sozialstaatsgebot 408
 Sozialversicherungssystem 375
 Spezialist 42
 Spezifität 124, 487
 – von 40% 269
 Sprunggelenksdistorsion 267
 Standard 394
 – medizinischer 394
 – rechtlicher 394
 Standard-Antidepressiva 335
 Standardlotterie 209
 STARD 492
 Stärke einer Empfehlung 234, 236
 Stimulation, therapeutische elektrische 326
 Strahlenbelastung reduzieren 270
 Stratifizierung 80
 Streuung, zufällige 487
 Strukturierung von Patientenerfahrungen 57
 Studie des Monats 440
 Studien 51, 62
 – chirurgische 467
 – diagnostische 492
 – klinische 63
 – ökonomische 68
 – qualitative 159
 – randomisierte 76, 332
 – randomisierte kontrollierte 137, 335
 – randomisierte kontrollierte klinische
 219, 443
 – retrospektive 136
 – symptomevaluierende 293
 Studienauswertung 491
 Studiendauer 273
 Studienergebnisse 39f.
 – verzerrte 405
 Studienforschung 467
 – klinische 468

Studienkritik 271
 Studienmethodik 273
 Studienprojekt 467
 Studienqualität 232
 Subgruppen 78
 Subgruppenanalysen 79
 Suche, sensitive oder spezifische 255
 Suchstrategien 91
 Supinations-Eversionsverletzung 267
 Surrogat-Endpunkte 147
 Surrogatparameter 37
 Symptom 93
 Symptomdefinition 96
 Symptom-Negative 97
 Syndesmosenverletzungen 267
 Systemebene 61, 275
 Systemscheidungen, Evidenzbasierte 61
 Systemversagen 407

T

Take home message 423
 Technologiebewertung 191
 Test
 – genetischer 119
 – statistischer 144, 489
 Testeigenschaften, Diagnostik 124
 Testen 485, 489
 – sequenzielles 129
 Testergebnisse
 – Patientenmanagement 130
 Theorie der Medizin 18
 Therapie
 – leitliniengerechte 177
 – medikamentöse 305
 Therapieentscheidungen 32
 Therapie-Evaluation 75
 Therapierichtungen, besondere 402
 Therapiestudien 415
 Training 428
 – repetitives sensomotorisches 326
 Train-the-Trainer-Seminare 441
 Triage, diagnostische 290
 Trias 61
 Trugschluss 77
 Tuberkulose-Prophylaxe, HIV-Infektion 150
 Tutor 422

U

Übereinstimmung
 – von Bewertern 153
 Übersicht, systematische
 – Ein- und Ausschlusskriterien 151

Übersichtsarbeit 339
 – systematische 219, 339, 443
 Übertragbarkeit 484
 Umsetzungshilfen 254
 Unfallversicherung 403
 Unmasking 98
 Unsicherheitsbereich 140
 Untergruppen 78
 Untergruppenanalyse 155
 Unterricht, gelungenerer 417
 Untersuchungs- und Behandlungsmethoden 403
 Unwirtschaftlichkeit 11
 Unwissenheit 89

V

Validierung 119
 Validität 40, 135, 340, 365, 491
 – externe 483
 – interne 483f.
 Validitätskriterien 415
 Verbesserungszyklus 287
 Verblindung 138, 272, 485
 Verbotsvorbehalt 385
 Verfahrensfehler 407
 Verfahrensordnung
 – G-BA 405
 Verfahrensordnung des Gemeinsamen Bundes-
 ausschusses 380
 Verhaltensänderungen 299
 Verification-Bias, diagnostische Tests 124
 Verlässlichkeit 177
 Verletzungen, knöcherne 267
 Vernetzung, praxisorientierte 475
 Versorgung
 – integrierte 286
 – koordinierte 284
 Versorgungsablauf 361, 365
 Versorgungsbereiche 45
 Versorgungskette 286
 Versorgungsqualität 361
 Versorgungsrealität 405
 Vertebroplastie 345
 Vertrauen 41
 Vertrauensbereich 487
 Vertrauensintervall 487
 Verzerrung 484, 486
 Verzicht auf Röntgenaufnahmen 267

Vorbehalt des Gesetzes 405
 Vorgehen
 – deduktives 262
 – diagnostisches 262
 – induktives 262
 – leitlinienorientiertes 341
 Vortest-Wahrscheinlichkeit 124, 292

W

Wahrscheinlichkeitsverhältnis
 – Nachteile 127
 – positiv und negativ 125
 – Testeigenschaften 124
 – Vorteile 128
 Weiter- und Fortbildung 421
 Weiterbildungsstrukturen für Patienten- und
 Verbrauchervertreter 434
 Wert
 – positiver prädiktiver 93
 – prädiktiver
 – positiv und negativ, Def. 127
 Werteentscheidung 61
 Wertvorstellungen 237
 Wirbelfrakturen, osteoporotische 345
 Wirksamkeit 275, 402
 – klinische 389
 Wirtschaftlichkeit 203, 402
 Wissen, bestes verfügbares 361
 Wissenschaft 75
 Wissenschaftseinrichtungen, studienbezogene 467
 Wissenschaftsmethodik 75
 Wissenschaftsprozess 77
 Wissensvermittlung 361
 Worst-case-Analyse 139

Z

Zeitaufwand 90
 Zeitschrift für ärztliche Fortbildung und Quali-
 tät im Gesundheitswesen 439
 Zertifizierung von Kursen 425
 Zielgrößen 361, 365
 Zielgrößenerfassung 110
 Zufall 486
 Zufallsfehler 134
 Zufriedenheit 37
 Zweifel 3

Die Schallmauer durchbrechen

Basis- und Fortgeschrittenenwissen Ultraschall

- Inhalt von Grund-, Aufbau und Abschlusskurs
- Kursbuch und diagnostischer Leitfaden
- Alle US-Bilder mit korrespondierendem Schema
- B-Bild, (Power-)Doppler, Echosignalverstärker

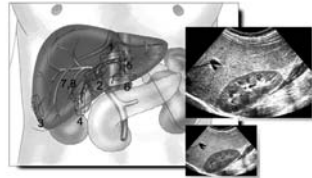
Organ für Organ in klaren, aufschlussreichen Bildern – für ein kursübergreifendes Arbeiten:

- Grundlegende sonographische Anatomie
- Häufige, leicht zu erhebende Befunde
- Seltene, schwerer erkennbare Krankheitsbilder

5. neu bearbeitete Auflage

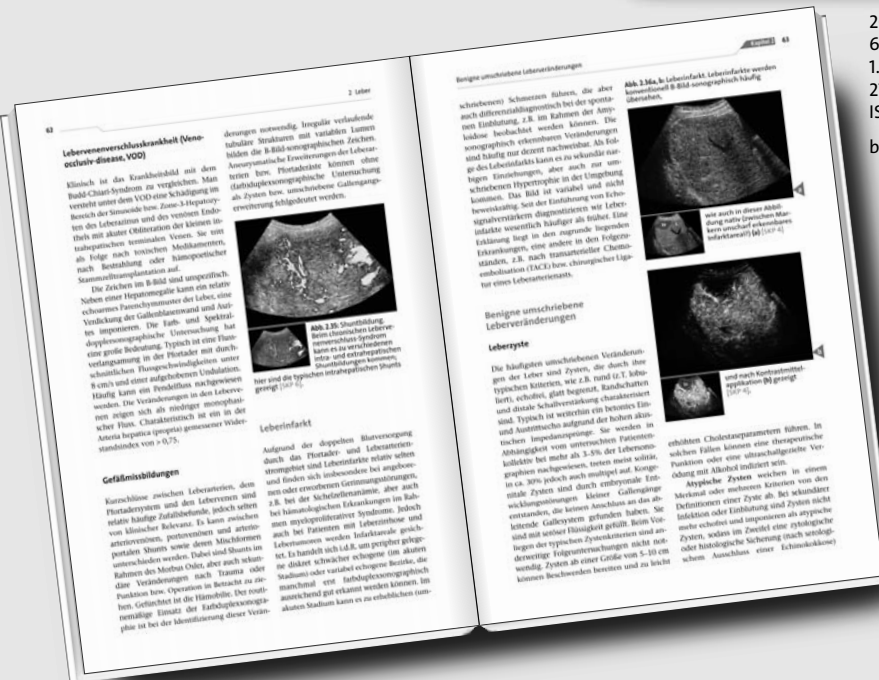
C. F. Dietrich Ultraschall-Kurs

Organbezogene Darstellung von
Grund-, Aufbau- und Abschlusskurs
nach den Richtlinien von KBV, DEGUM,
ÖGUM und SGUM



Deutscher
Ärzte-Verlag

2006, 450 Seiten,
677 Abbildungen in
1.421 Einzeldarstellungen,
27 Tabellen
ISBN 978-3-7691-0451-6
broschiert € 59,95



Bestellungen bitte an Ihre Buchhandlung oder Deutscher Ärzte-Verlag, Versandbuchhandlung:
Postfach 400244, 50832 Köln. Tel. (0 22 34) 7011-314 / Fax 7011-476 / E-Mail: vsbh@arztverlag.de

Mehr Information: arztverlag.de

Deutscher
Ärzte-Verlag

Irrtümer und Preisänderungen vorbehalten. Preis zzgl. Versandkosten € 4,50
Deutscher Ärzte-Verlag GmbH - Sitz Köln - HRB 106 Amtsgericht Köln

Innere Krankheiten sporttherapeutisch behandeln

Didaktisch kompetent vermittelt das Standardwerk sowohl Basiswissen als auch sporttherapeutische Konsequenzen.

- Atemwegserkrankungen
- Herz-Kreislauf-Erkrankungen
- Erkrankungen des Blutes (Leukämie, AIDS etc.), der Gefäße, der Verdauungsorgane und der Niere
- Krebserkrankungen
- Rheumatologische Erkrankungen
- Rehabilitation

Ihr Know-how für eine interdisziplinäre Sport- und Bewegungstherapie.



3. überarbeitete und erweiterte Aufl. 2005,
479 Seiten, 114 Abbildungen, 52 Tabellen
ISBN 978-3-7691-0411-0

broschiert € **49,95**

Bestellungen bitte an Ihre Buchhandlung oder Deutscher Ärzte-Verlag, Versandbuchhandlung:
Postfach 400244, 50832 Köln. Tel. (0 22 34) 7011-314 / Fax 7011-476 / E-Mail: vsbh@aerzteverlag.de

*Irrtümer und Preisänderungen vorbehalten. Preise zzgl. Versandkosten € 4,50
Deutscher Ärzte-Verlag GmbH - Sitz Köln - HRB 106 Amtsgericht Köln*

Grundlagen der EbM – Schritt für Schritt in Fallbeispielen

Läuft Ihnen der medizinische Fortschritt manchmal davon? Wie filtern Sie aus der Fülle neuer Studienergebnisse das **richtige und für Sie wichtige Wissen** heraus?

Neue Erkenntnisse zu Diagnostik und Therapie werden fast nur noch durch groß angelegte Studien gewonnen, die jede für sich Möglichkeiten des Irrtums, der Verzerrung und der Manipulation aufweisen. Wert oder Unwert solcher Studien erschließen sich fast nur noch dem Experten. Abhilfe schafft der (selbst-)kritische Denkansatz der **Evidenzbasierten Medizin**. Diese hilft Ihnen, frühzeitig zu erkennen, was Ihnen zur **richtigen Entscheidung** und dem Patienten zum Wohle dient.

Dieses Lehrbuch zeigt Ihnen Schritt für Schritt anhand **typischer Problemstellungen** einzelner Fachgebiete wie Sie den „Dschungel“ des medizinischen Know-hows durchforsten. So können Sie selbst differenzieren, ob ein neues, angepriesenes Verfahren Marketing oder tatsächlich Fortschritt bedeutet.

EbM ist keineswegs die „Lehre von der doppelblind-randomisierten Studie“, sondern verhilft Ihnen auch bei komplexer Studienlage zu einer transparenten und souveränen Entscheidung.

Mit EbM bleiben Sie dem medizinischen Fortschritt auf den Fersen!

- ▲ Der Bedarf für EbM
- ▲ Basiswissen zur kritischen Bewertung wissenschaftlicher Informationen
- ▲ Praktische Umsetzung von EbM im ärztlichen Alltag
- ▲ EbM in den Strukturen der Gesundheitsversorgung
- ▲ EbM in Aus-, Weiter- und Fortbildung
- ▲ Deutsche EbM-Organisationen



Prof. Dr. med. Regina Kunz
Stellvertretende
Institutsleitung
Institut für klinische
Epidemiologie,
Universitätsspital Basel



Prof. Dr. rer. nat. Dr. med. Günter Ollenschläger
Leiter des Ärztlichen
Zentrums für Qualität in
der Medizin, Berlin



Prof. Dr. med. Dr. phil. Heiner Raspe
Direktor des Instituts für
Sozialmedizin der
Medizinischen Universität
Lübeck



Dr. med. Günther Jonitz
Facharzt für Chirurgie,
Präsident der Ärztekammer
Berlin



Prof. Dr. med. Norbert Donner-Banzhoff
Leiter des Studienprogramms
„Klinische Evaluation“, Abtei-
lung für Allgemeinmedizin,
Rehabilitative und Präventive
Medizin, Philipps-Universität
Marburg

In Kooperation mit
dem Deutschen
Netzwerk
Evidenzbasierte
Medizin e. V.



**Deutscher
Ärzte-Verlag**

aerzteverlag.de

ISBN 978-3-7691-0538-4



9 783769 110538 4