

6 Evidenz-basierte Medizin und Leitlinien in der Kardiologie. Fördern sie Wissenschaft oder bürokratische Strangulierung?¹

Frank Praetorius

Eine Medizin, die „Evidence-based“ sein will, betrachtet solche Grundlegung gleichermaßen als Forschungs- und Anwendungsprojekt – und verdoppelt auf diese Weise auch die Probleme im ethischen Bereich. Auf dem wissenschaftlichen Sektor will die *Deutsche Forschungsgemeinschaft (DFG)* mit Hilfe der Evidenzkriterien eine Qualitätssteigerung der so genannten *patientenorientierten klinischen Forschung* erreichen, womit vor allem die *Klinischen Studien* gemeint sind (DFG, 2002)⁽¹³⁾. Diese Forschungsrichtung ist ein Stiefkind der gegenwärtigen akademischen Strukturen (Raspe 2001)⁽³⁷⁾. Deshalb fordert die DFG gleichzeitig eine Beseitigung der strukturellen Defizite an den Hochschulkliniken und einen Abbau der weit verbreiteten *"Pro forma"-Forschung*.

In der klinischen Praxis steht die Entwicklung und Implementierung evidenzbasierter Leitlinien im Vordergrund. Anwendungsprobleme entstehen hier besonders beim Versuch der Abgrenzung oder Versöhnung von wissenschaftlicher und ökonomischer Rationalität (Davidoff et al. 2001; Schulman et al.)^(9; 40). Generell möchte die Evidenzbasierte Medizin (EBM) den Prozess der ärztlichen Professionalisierung „in Richtung zunehmender wissenschaftlicher Rationalität verändern“ (Vogd 2002)⁽⁴⁷⁾. Die zugrunde liegende Intention ist es, die Erkenntnisse der Basiswissenschaft "Klinische Epidemiologie" in ein praxistaugliches und anwendernahes Instrumentarium zu verwandeln (Perleth and Raspe 1998)⁽³¹⁾. Allerdings kann eine ideologisch überzogene Verwissenschaftlichung über Schematisierung und Bürokratisierung des ärztlichen Alltags in ihr Gegenteil umschlagen: In Deprofessionalisierung durch Überregularisierung (Praetorius and Sahn 2001)⁽³⁴⁾. Für die Diskussion dieser Fragen dürfte die Kardiologie als das Gebiet mit der vermutlich größten Anzahl bestehender Richtlinien beispielhaft geeignet sein, besonders weil randomisierte kontrollierte Studien (RCT) bei kardiovaskulären Erkrankungen traditionell stärker unter dem Einfluss von „steering and data monitoring committees“ stehen als bei den meisten anderen Erkrankungen (DeMets and Califf 2002)⁽¹¹⁾.

Evidenzbasierte Kardiologie

Bei den kardiovaskulären Krankheiten (CVD = Cardiovascular Diseases) wurden durch die Einführung und gegenseitige Abgrenzung von neuen Substanzen und therapeutischen Techniken seit etwa 1980 zunehmend Studien mit hohen Fallzahlen (darunter Megastudien mit über 10.000 Patienten) erforderlich, deutlich mehr als auf dem diagnostischen Sektor (DeMets and Califf 2002)⁽¹¹⁾. Ehe die Ergebnisse in die Leitlinien von nationalen Gesellschaften eingehen, wird gewöhnlich eine Anzahl von Studien einer so genannten Metaanalyse unterzogen (deren wissenschaftlicher Wert hier nicht diskutiert werden soll). Die Entwicklung führt tendenziell zu einer Internationalisierung der Leitlinien; diese werden zwar in der Regel noch auf nationaler Ebene formuliert, unterscheiden sich aber inhaltlich wenig, bis hin zur grafischen Gestaltung (Von Schacky 2002)⁽⁴⁸⁾. In Deutschland wird angestrebt, auf der Basis des Leitlinien-Clearance-Verfahrens (welches die alte Konsensmethode zunehmend ablöst) erstmals rechtlich verbindliche Leitlinien zu entwickeln (Dierks 2003)⁽¹⁴⁾.

¹ Evidence-based Medicine and Clinical Guidelines in Cardiology. Promoting Science, Practice, or Bureaucracy? In: Meulen RHJ ter, Biller-Andorno N, Lenk C, Lie R (Eds.) (2005) Evidence-based Practice in Medicine and Health Care. A Discussion of the Ethical Issues. Springer, Berlin-Heidelberg

Leitlinienanbieter

Eine Zählung aller weltweit zur Verfügung stehenden Leitlinien ist kaum noch möglich; man kann von über 1.000 kardiologischen Leitlinien ausgehen. Dementsprechend sehen sich Kliniken und Spezialpraxen zur Auswahl geeigneter Leitlinien nach eigenen Kriterien genötigt. Als Beispiel kann die von der Medizinischen Klinik I in Offenbach am Main (Chefarzt Prof. Dr. H. Klepzig; <http://www.klinikum-offenbach.de/med1/>) für die Klinik und für die einweisenden Ärzte ins Internet gestellte Leitlinienliste gelten.

Die „Deutsche Gesellschaft für Kardiologie (DGK)“ listet derzeit im Internet 38 so genannte Leitlinien, wozu allerdings Positionspapiere, Kommentare, Berichte und Stellungnahmen gerechnet werden (lediglich 11 Leitlinien im eigentlichen Sinn). Die folgende Liste der internationalen Leitlinienanbieter wird von der „Ärztlichen Zentralstelle Qualitätssicherung“ angeboten (► *Tabelle 6-1*):

Table 6-1
International Providers of Guidelines in Cardiology (8 January, 2004);

<http://www.leitlinien.de/leitlinienanbieter/fachspezifisch/>

ACC (American College of Cardiology)	USA	English
AHA (American Heart Association)	USA	English
American Healthways	USA	English
BCS (British Cardiac Society)	Great Britain	English
Canadian Cardiovascular Society	Canada	English
Cardiac Society of Australia and New Zealand	Australia	English
Deutsche Gesellschaft für Kardiologie	Germany	German
European Society of Cardiology	Europe	English
Heart Failure Society of America	USA	English
Heartfoundation of Australia	Australia	English
Heartfoundation of New Zealand	New Zealand	English
Nederlandse Vereniging voor Cardiologie	Netherlands	Engl., nl
NHLBI Cardiovascular Information (The National Heart, Lung, and Blood Institute)	USA	English
NHLBI Clinical Guidelines	USA	English
Schweizerische Gesellschaft für Kardiologie	Schweiz	Germ., French
Société Française de Cardiologie	France	French
Österreichische Kardiologische Gesellschaft	Austria	German

Die Qualität von Leitlinien

In der internationalen Kardiologie gelten die in Zusammenarbeit des American College of Cardiology (ACC) und der American Heart Association (AHA) entstandenen Leitlinien als vorbildlich. Dort werden die Empfehlungsgrade und Evidenzklassen (Class I, II, III und Grade A, B, C) genannt, nach denen die Leitlinien klassifiziert wurden (► *Tabelle 6-2*):

Als Beispiel dient eine aktuelle Stellungnahme und Empfehlung der AHA und der Centers for Disease Control and Prevention (CDC) (Pearson et al. 2003)⁽²⁹⁾, in der es um die Evidenz der Beziehung zwischen Markern der Entzündung und arteriosklerotischen Herzkrankheiten geht, vor allem um den verbesserten Marker hs-CRP (für high-sensitivity C-reactive protein),

der für die Diagnose von „entzündlichen“ Exazerbationen der Koronararteriosklerose als derzeit am besten geeignet gilt. Die für die Empfehlungen der AHA verwendeten Kriterien bei Schlussfolgerungen (im Sinne „medizinischer“ Kausalität) werden ebenso dargelegt wie die der Evidenzqualität von neuen gedanklichen Verknüpfungen (Assoziationen).

■ **Table 6-2**

American College of Cardiology/American Heart Association classification of recommendations and levels of evidence

Classification of Recommendations	Levels of Evidence
<p><u>Class I</u> Conditions for which there is evidence and/or general agreement that a given procedure is useful and effective.</p> <p><u>Class II</u> Conditions for which there is conflicting evidence and/or divergence of opinion about the usefulness/efficacy of a procedure or treatment.</p> <p><u>Class IIa</u> Weight of evidence/opinion is in favor of usefulness/efficacy.</p> <p><u>Class IIb</u> Usefulness/efficacy is less well established by evidence/opinion.</p> <p><u>Class III</u> Conditions for which there is evidence and/or general agreement that the procedure/treatment is not useful/effective, and in some cases may be harmful.</p>	<p><u>Level of Evidence A</u> Data derived from multiple randomised clinical trials.</p> <p><u>Level of Evidence B</u> Data derived from a single randomised trial or nonrandomized studies.</p> <p><u>Level of Evidence C</u> Consensus opinion of experts</p>

Es lohnt sich, die kritische Stellungnahme für die Anwendung der hochsensitiven CRP-Bestimmung gegenüber anderen Entzündungsparametern und für die Einschränkung auf präzise definierte Risikogruppen im Einzelnen zu betrachten: Sorgfältig wird zwischen Patienten mit bekannter koronarer Herzkrankheit oder einem durch andere Parameter belegtem „hohem koronarem Risiko“ und solchen ohne Hinweise auf eine Herzerkrankung unterschieden.⁽²⁹⁾ Auf die noch unklaren therapeutischen Konsequenzen wird deutlich hingewiesen und die Anwendung in der Sekundärprävention abgelehnt. Die Versuchung, sich mit Hilfe der hs-CRP-Bestimmung die Entscheidung zu aggressiver medikamentöser Behandlung – etwa durch frühzeitigere Eingriffe in das Gerinnungssystem – leichter zu machen, wird durch diese Empfehlungen nicht gestützt.

Die Arbeit enthält eine nicht weiter ausgeführte ethische Aussage: Evidenzbasierte therapeutische Entscheidungen dürfen niemals von einem Test abhängig gemacht werden, der im Hinblick auf therapeutische „Benefits“ noch ungeklärt ist, auch wenn er im Hinblick auf das aktuelle pathophysiologische Denken viel versprechend zu sein scheint. Leider werden neue diagnostische Methoden in der Kardiologie nur selten mit dem Ziel untersucht, ihre Effizienz im Hinblick auf den klinischen Verlauf und auf künftige therapeutische Konsequenzen methodisch abzuklären.

Unterschiedliche Länder

Beim Versuch, die Angemessenheit der diagnostischen Koronarangiografie bei einer Vielzahl von möglichen Indikationen zu bestimmen, fällt ein Mangel an methodisch einwandfreien Studien auf. An Stelle kontrollierter Studien findet sich nur ein Expertenkonsens, der in einer Grauzone wissenschaftlicher Erkenntnis vermitteln muss (Perleth and Raspe 1998)⁽³¹⁾, und bedauerlicherweise nur über ein *Level of Evidence C* mit einer Empfehlungsstärke von *Class IIb* verfügt. Dennoch wurde die Methode zum „Gold-Standard“. Das Problem wurde in der Diskussion im *LANCET 2002* deutlich (Dissman and De Ridder 2002)⁽⁴⁵⁾, bei der es um die große Zahl von Koronarangiografien in Deutschland (als europäischer Spitze) ging. Befürwortern und Gegnern der Kritik im *LANCET* wurde entgegengehalten: *They do not discuss the effect of increased use of invasive procedures on*

morbidity and mortality from cardiovascular disease (Baldeweg, siehe (15)). Statt sich über die Zweckmäßigkeit der Indikation für den Patienten auseinanderzusetzen, werden kurzschlüssig Untersuchungszahlen und Kosten assoziiert. Das ähnelt dem Vorgehen von Politikern – und so nimmt es nicht Wunder, dass die ökonomisch aufwendige Herzkatheteruntersuchung politisch im Zentrum der Kritik an „unwirtschaftlicher Medizin“ bleibt.

Die Richtlinien und Empfehlungen der Deutsche Gesellschaft für Kardiologie (DGK) werden in Zusammenarbeit mit der „Arbeitsgemeinschaft der wissenschaftlichen medizinischen Fachgesellschaften (AWMF)“ herausgegeben (<http://www.uni-duesseldorf.de/www/awmf>). Schon 2002 wurde von *Lauterbach* kritisch gesehen (Lauterbach and Redaelli 2002)⁽²⁶⁾, dass der wichtigste Mangel dieser und anderer europäischer Leitlinien die fehlende Evidenzklassifizierung ist. Bei fast allen der über 100 im Jahre 2000 ausgewerteten Leitlinien fehlten die Evidenzklassenangaben der Literatur – mit dem entsprechenden Mangel an Transparenz als Folge. Außerdem waren anwenderfreundliche Algorithmen mit dem Ziel einer besseren Praktikabilität der Leitlinien selten. Sie werden auch heute weitgehend vermisst, obgleich deutschen Leitlinien „eine kurze, sprachlich eindeutige Zusammenfassung“ (Kunz et al. 2000)⁽²⁵⁾ haben sollten, so dass sie zum schnellen Nachschlagen in einer Entscheidungssituation dienen könnten.

Auch bei den 2001 veröffentlichten „Leitlinien zur Therapie der chronischen Herzinsuffizienz“ der DGK (Erdmann and Hoppe 2001)⁽¹⁷⁾ werden an einem entscheidenden Punkt die *Levels of Evidence* deutlich anders definiert als von AHA/ACC. Unter „B“ heißt es dort: „Positive Aussage gestützt durch eine randomisierte kontrollierte klinische Studie oder durch klinische Erfahrung.“ Wessen „klinische Erfahrung“? Dieser Qualität fehlt jede präzise Definition. Sie kommt in den AHA/ACC-Empfehlungen nicht vor und könnte allenfalls unter „Evidence C“, *Consensus opinion of experts*, subsumiert werden. Der als „D“ hinzugefügte Grad („Negative Aussage gestützt durch eine oder mehrer klinische Studien“) bedeutet in den Einteilungen der *Canadian* und *US-Task Force* etwas anderes und bezieht sich dort ausschließlich auf Leitlinien zur Gesundheitsuntersuchung (siehe bei Kunz a.a.O.). Es ist verständlich, wenn ein kritischer Haus- oder Facharzt deshalb zögert, dieser sonst gut formulierten Leitlinie zu folgen.

Zur Kritik der EBM.

Die Rolle von kontrollierten prospektiven Registern.

Einige Autoren argumentieren vehement gegen den Begriff „EBM“. Aber Worte kann man immer durch andere ersetzen, beispielsweise EBM durch „research enhanced health“, was nach *R. B. Haynes* (Haynes et al. 2002)⁽²¹⁾ die heutige Auffassung von EBM etwas besser abbildet. Klinische Überlegungen sollen nach *Haynes* an erster Stelle die aktuelle klinische Situation und den Status präsens erfassen, diese mit Hilfe evidenter Forschungsergebnisse modifizieren (zweitens) und drittens die Wünsche des Patienten und seine vermuteten Reaktionen einbeziehen, ehe schließlich die so genannte *Clinical expertise* alle diese Überlegungen zusammenfasst und jetzt dem Patienten eine Empfehlung gibt. Dem Patienten Evidenz in einer Form anzubieten, die ihm eine intelligente, „informierte“ Entscheidung ermöglicht, ist eine Herausforderung, die durchaus ebenfalls einer evidenzbasierten Forschung bedarf (Haynes a.a.O.).

Das leitet zu jener mehr philosophischen Kritik über, die in der deutschen Debatte (Rogler and Schölmerich 2000; Raspe 2001)^(38;37) über *Evans' evidence-based* versus „*Evidence-Biased Medizin*“ – oder: *Die trügerische Sicherheit der Evidenz*“ im Vordergrund stand. Dies wäre eine wichtige Diskussion, wenn die EBM einen Anspruch als Disziplin der Philosophie oder der Wissenschaftstheorie erheben würde. Obgleich es zutreffen mag, dass einige EBM-Ideologen das glauben und einige Politiker ihnen darin allzu willig folgen, sieht die Mehrzahl der Kliniker und EBM-Spezialisten die EBM als eine bloße Hilfsmethode für die Praxis an: Sie soll die Anwendung von mehr Evidenz bei ärztlichen Entscheidungen durch Anwendung von Ergebnissen der Klinischen Epidemiologie fördern.

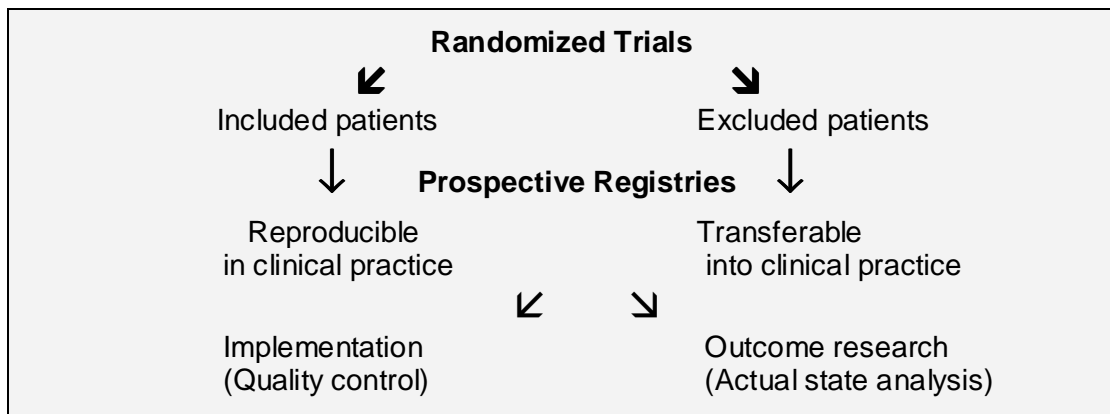
Bei Studien ausgeschlossene Patienten: Eine der Zielgruppen für Register

„Randomised controlled Trials“ (RCT) sind zum entscheidenden Hilfsmittel der Forschung geworden, wenn es darum geht, Evidenz zu entwickeln. RCT sind ohne eine genaue Abgrenzung nicht korrekt durchzuführen. Vor allem der Ausschluss älterer oder jüngerer Patienten, die Einschränkung auf Männer oder Frauen, auf Rassen oder auch auf spezielle klinische Situationen (wer wird behandelt, wer nicht?) ist häufig nötig, ebenso wie die Eingrenzung durch die festzulegenden diagnostischen Kriterien der Krankheit und durch Laboruntersuchungen. Das Problem, wie man unterrepräsentierte Minoritäten adäquat in klinische Studien einschließt, scheint unüberwindbar zu sein. Darüber hinaus können neue Labortests dazu führen, dass eine schon durch RCTs untersuchte Krankheit neu klassifiziert wird und damit neue Patientengruppen in RCT ein- oder auszuschließen sind. So haben zum Beispiel in der aktuellen Kardiologie neue Labortests eine völlig neue Differenzierung zwischen Myokardinfarkten mit oder ohne ST-Hebung ermöglicht (siehe unten). Und natürlich bleiben die Standardfälle für eine Exklusion aus Studien und damit aus der kontrollierten Forschung bestehen: Multimorbidität, Schock, Zustand nach Reanimation etc..

Aus diesen Gründen sind die in RCT untersuchten Patienten oft nicht mit jenen identisch, die im Klinikalltag selbstverständliche Behandlungsfälle darstellen. Der Anteil der „eingeschlossenen“ Patienten wird auf *selten mehr als 50%* des Gesamtkrankenguts geschätzt (Kunz et al. 2000)⁽²⁵⁾. In der EBM-Literatur wird deshalb definitorisch die „Efficacy“ in randomisierten kontrollierten Studien (RCT) von der „Effectiveness“ in der Versorgungsforschung getrennt: Der Begriff der *Efficacy* bezieht sich auf die Wirksamkeit unter idealen Studienbedingungen im Labor oder in einer RCT, während *Effectiveness* in der „realen Welt“ des Krankenhauses oder der Arztpraxis gilt, also für das Feld der Versorgungsforschung (Kunz a.a.O., Haynes a.a.O.).* Der Kardiologe *J. Senges* hat die gegenseitige Ergänzung evidenzgestützter Leitlinien einerseits und von prospektiven Registern andererseits hervorgehoben (Senges and Gitt 2002)⁽⁴¹⁾. Register entsprechen nach *Senges* dem „Evidence level B“ unter der Bedingung, dass sie wie das ALKK-Register (siehe Praetorius 1999⁽³⁵⁾) „prospektiv“ sind, d. h. die Aufnahme von Patienten erfolgt nach vorher festgelegten und kontrollierten Parametern wie in RCTs – und nach dem „intention-to-treat“-Prinzip. Ihre praktische Stärke liegt darin, dass sie im Gegensatz zu dem stark selektierten Patientengut der kontrollierten Studien das gesamte Spektrum der Bevölkerung erfassen. So wird die Kontrolle der Reproduzierbarkeit von RCT und deren Übertragbarkeit auf die in den RCT ausgeschlossenen Patientengruppen ermöglicht (▶ *Abbildung 6-1*).

■ Figure 6-1

Randomized Trials and Prospective Registries modified from J. Senges and A.K. Gitt, 2002



* Nach der Darstellung von Jürgen Windeler u. Gerd Antes gibt es keine "scharfe" allgemein gültige Definition

Unter den akuten koronaren Syndromen fand sich im Register der deutschen kommunalen Krankenhäuser (ALKK) eine unerwartete Dimension der so genannten Nicht-ST-Hebungsinfarkte. Die Diagnose von Myokardinfarkten (MI) ohne (englisch None = N) ST-Elevation (STE) im EKG, international akronymisiert NSTEMI, ist durch die neue Standardmethode der Troponinbestimmung (anstelle der CK-MB) zu sichern. Im ALKK-Register machten die NSTEMI-Herzinfarkte unter 9.947 Infarkten 41% aus (STEMI 46%, instabile Angina pectoris 13%), das sind etwa doppelt so viele wie früher angenommen (Senges and Gitt 2002)⁽⁴¹⁾. Der Ist-Zustand hat sich also geändert, und andere, frühzeitig einsetzende invasive Behandlungsstrategien wie Glycoprotein IIb/IIIa-Antagonisten und „low-molecular-weight“ Heparine sind denkbar (Bertrand et al. 2002)⁽⁴⁾, wenn auch noch nicht evidenzbasiert. Ohne ein Register wäre diese Tatsache wesentlich später aufgefallen.

Ein wichtiges Ergebnis ergab auch die prospektive Beobachtungsstudie der deutschen Herzchirurgen (COCPIT Study Group (Deng et al. 2000))⁽¹²⁾ über die Erfolge der Herztransplantation: Nach vielen Jahren „sicheren Wissens“ (seit dem Durchbruch durch Cyclosporin 1980, aus ethischen Gründen ohne RCT) fand sich überraschend nur bei den schwersten Fällen mit hohem, nicht bei mittlerem oder geringem Todesrisiko ein positiver Effekt der Transplantation für die Patienten (der leider auch nur für die ersten 5 Monate nach der Transplantation nachweisbar ist), so dass jetzt international eine RCT gefordert wird.

Die Teilnahme an Registern setzt bei den aktiven Ärzten die Bereitschaft zu erheblichem bürokratischem Aufwand und zu Kontrollen und Überprüfungen voraus. Dies kann durch den auf diese Weise möglichen Vergleich des eigenen Tuns mit der Gesamtgruppe (Praetorius 1999)⁽³⁵⁾ und durch den wissenschaftlichen Gewinn subjektiv kompensiert werden. Allerdings muss auch bei Registern mit dem Anspruch von Qualitätssicherungsprogrammen auf Interferenzen mit den Interessen von Herstellern geachtet werden, welche die Implementierung und Verbreitung ihrer Präparate durch die Finanzierung von Registern fördern wollen. So im Falle des „Qualitätsregisters ERICA“ (Establishing Risk Reduction in Congestive Heart Failure through Add on Therapy), in dem es um die Implementierung der Therapie mit β -Rezeptorenblockern bei der Herzinsuffizienz geht. Das Medikament *Bisoprolol* wird dort nicht genannt, die übliche Autorenerklärung liegt vor (Antoni et al. 2002)⁽²⁾. Die zugehörige Softwarefirma *mediNET* benennt in ihrer Internet-Information die Quelle („Qualitätsregister ERICA, in Kooperation mit MERCK“) und die „Fachinformation Concor® COR“. Diese Bemerkung soll keinen moralischen Vorwurf darstellen, weil konkrete Verdachtsmomente fehlen. Aber da Transparenz die Basis der angewandten Ethik in der Klinik ist, kann man feststellen, dass eine Nennung der Verbindung in der Autorenerklärung jede Kritik vermieden hätte.

RCT und Guidelines – sind sie wirklich up-to-date?

Große randomisierte und multizentrische Studien benötigen eine systematische und langwierige Planung und – nach Abschluss der Studie – einen oft mehrjährigen Zeitraum für die wissenschaftliche Auswertung und angemessene Publikation. Man schätzt im Allgemeinen, dass mindestens zwei, oft bis zu fünf Jahre erforderlich sind. Es entsteht daraus eine ethische Frage: Kann die Anwendung einer neuen Therapie, die nach kleineren Studien beispielsweise schon Level B, Class IIb der Tabelle 2 erfüllt, allein vom Vorliegen großzahliger RCT abhängig gemacht, also dem Patienten solange vorenthalten werden? Ideal wäre es, wenn jede Therapie auf den typischen „benefits“ der großen Studien beruhen könnte, auf Überlebenszeit, Vermeidung bestimmter klinischer Ereignisse (Myokardinfarkt, Schlaganfall etc.) und verbesserter Lebensqualität – zu „vernünftigen“ Kosten (DeMets and Califf 2002)⁽¹⁰⁾. Um das Problem zu lösen, werden so genannte Surrogat-Parameter benutzt, um raschere Ergebnisse mit weniger Patienten und Kosten zu erzielen. Das ist in frühen Phasen der Forschung auch deshalb sinnvoll, weil bei der Entwicklung neuer Therapien durchaus deduktive pathophysiologische Gesichtspunkte von primärer Bedeutung sein können (DeMets and Califf 2002)⁽¹⁰⁾. Aber Surrogat-Endpunkte von Studien können zu schweren Irrtümern führen; beispielsweise garantiert das Surrogat „erfolgreiche Blutdruck-

senkung“ bei Hypertonikern nicht den gleichen Erfolg wie bei den wesentlichen Parametern Tod und Myokardinfarkt als „Outcome“ (ALLHAT 2002)⁽¹⁾.

Nur 30 % der heute verfügbaren Therapien beruhen auf einem höheren Evidenzniveau. Bei neuen Medikamenten oder technischen Behandlungskonzepten gibt es eine Tendenz, diese Therapien auf der schwächeren Basis des Expertenkonsenses einzuführen, der rascher herzustellen ist (Level C). In der Kardiologie galt das für eine längere Zeit bei der Ballondilatation (PTCA) und später bei Stents. Deshalb wurde in Deutschland das erste prospektive Register der „Arbeitsgemeinschaft leitender Kardiologischer Krankenhausärzte (ALKK)“ in Deutschland begründet – zunächst mit dem Ziel, einen kritischen Überblick über das eigene Tun zu bekommen. Bis 1999 waren über 200.000 PTCA erfasst (Praetorius 1999)⁽³⁵⁾; seit 2001 ist das Register ein Teil gesetzlich vorgeschriebenen Qualitätssicherung. Bei den regelmäßigen Vergleichen der eigenen Klinik mit dem Gesamtregister stellte sich ein überraschend einheitliches Verhalten der knapp 100 Kliniken heraus. Die Mehrzahl der Teilnehmenden richtete sich offenbar nach den gleichen wissenschaftlichen Arbeiten und passte die eigene der jeweils allgemeinen Vorgehensweise an. Diese geschlossene Gruppe konnte auf der Basis der vorhandenen Registerstruktur weitere Register durchführen und publizieren wie *“Determinants of mortality after cardiac surgery: results of the Registry of the Arbeitsgemeinschaft Leitender Kardiologischer Krankenhausärzte (ALKK) on 10525 patients”* (2000).

Medikamenteninteraktion und Interferenz von Leitlinien: Das zentrale Thema der Hausärzte

Zur Frage der Übersetzung von Forschungsergebnissen in die Versorgung von individuellen Patienten gibt es nur wenige wissenschaftliche Untersuchungen. RCTs konzentrieren sich häufig auf die Untersuchung eines einzelnen Medikaments. Bei gleichzeitiger Verordnung von mehreren Medikamenten können die möglichen Effekte grundsätzlich sowohl additiv wie synergistisch bzw. subadditiv, neutral oder negativ sein (Califf and DeMets 2002)⁽⁵⁾. Ähnliches gilt für die Interferenz von mehreren Leitlinien im gleichen klinischen Bereich. Bei Klinikern und Praktikern ist ein Gefühl der Verunsicherung die Folge - und sie zögern, die evidenzbasierten Leitlinien bei allen Patienten mit derselben Erkrankung anzuwenden.

Als Beispiel wird die ungeklärte praktische Bedeutung der Interaktion von Acetylsalicylsäure (ASS) und ACE-Hemmstoffen (ACE=Angiotensin Converting Enzyme) genannt. Beide Medikamente sind durch RCTs und Metaanalysen bestätigt und allgemein akzeptiert. Sie haben bei den gleichen Krankheiten unterschiedliche Zielpunkte: Gerinnung und Entzündung die eine, Blutdruck und Herzleistung die andere. Obgleich die Hinweise auf eine Abschwächung der ACE-Hemmer-Wirkung durch ASS seit mindestens 1998 deutlich sind, muss man sich zu diesem wichtigen Punkt auf die retrospektive Auswertung von Studien stützen (Califf and DeMets 2002)⁽⁵⁾, da eine RCT für genau dieses Problem nicht vorliegt (Joost 2003)⁽²²⁾.

Generell werden Wechselwirkungen von Medikamenten selten prospektiv im Hinblick auf bestimmte Indikationen und Dosierungen untersucht. Für die Hausärzte bleiben Fragen wie „hebt ASS die Wirkung von ACE-Hemmern auf und in welchem Ausmaß“, und welche Leitlinien soll man sich sozusagen „implementieren“ lassen? Es ist verständlich, wenn manche Ärzte ihren Patienten einfach erklären, dass ein oder zwei Tabletten besser verträglich sind als sieben oder mehr am Tage (► *Tabelle 6-3*), obgleich die evidenzbasierten Leitlinien die Kombinationen ausdrücklich empfehlen. Unter solchen Bedingungen leidet der Grundsatz des Informed Consent: Das ökonomische Motiv der Budgetschonung verschwindet hinter zweifelhaften medizinischen Begründungen (Kuhlmann 1999)⁽²³⁾. Aber es muss betont werden, dass der Ursprung dieses ethischen Konflikts in der EBM selbst zu suchen ist, die bisher nicht in der Lage war, die multiplen Forschungsergebnisse adäquat in die Welt der Klinik und Praxis zu übertragen.

Bei der Interaktion von Medikamenten oder Leitlinien von handelt es sich um ein vielschichtiges ökonomisches und ethisches Problem, dessen Lösung man nicht einfach der Hausarztmedizin und den Kliniken vor Ort anlasten kann. „Implementation“ kann nicht ein-

fach aus besserer Fortbildung und gut inszenierter Propaganda bestehen. Vielmehr sollten die EBM-Spezialisten generell evidenzbasierte Leitlinien für den Umgang mit Leitlinien erstellen, und ebenso die Forderung nach prospektiven Studien über Interaktionen und Interferenzen empfohlener Therapien unterstützen. Das gilt auch für die dringend notwendigen direkten Vergleichsstudien zur Elimination wenig nützlicher Therapien. Sicherlich ist es schwierig, für so etwas Sponsoren zu bekommen – aber es ist ein dringliche Aufgabe für die universitäre Forschung, zumal die Manager der EBM dazu neigen, solche Probleme unter dem Nebelschleier der „Implementationsdefizite“ als einziger Lösung vergessen zu machen.

Um die Schwierigkeiten zwischen EBM und der realen Welt der Hausärzte deutlich zu machen, soll die Praxisrealität am Beispiel von Hypertonie und Herzinsuffizienz in Tabelle 6-3 simuliert werden. Nach RCT werden in Leitlinien mindestens 7 Therapieformen mit hohem Evidenzniveau empfohlen. Der Hausarzt muss entscheiden, welche Therapie(n) er wählt. Er ist dabei in einer anderen Situation als der Spezialist und Klinikarzt, der in Substanzgruppen und Generic Names denkt und Kosten und Compliance weniger beachtet. Für die Darstellung in Tabelle 6-3 wurden für die Gruppen der ACE-Inhibitors (Enalapril), Betazeptorenblocker (Metoprolol), Calciumantagonisten (Amlodipin), Angiotensin-II-Rezeptor-Antagonisten (Losartan), Digitoxin, Chlorthalidon und Spironolactone Fantasienamen gewählt, die zugleich auf gewisse Effekte und aktuelle Diskussionen bei den Medikamenten hinweisen. Sie könnten problemlos durch Firmennamen ersetzt werden. Das wichtige Problem der Klasseneffekte von Medikamentengruppen mit identischen biologischen Zielen kann hier nur erwähnt werden. Besonders in der Gruppe der ACE-Hemmer werden Klinik und Praxis in der Verlegenheit gelassen, ob sie entweder an den Klasseneffekt glauben oder nur den einen RCT-getesteten ACE-Hemmer benutzen dürfen (Califf a.a.O.)⁽⁵⁾.²

■ Table 6-3

Hypertension and Heart failure: Choice of the Medicament(s)			Family doctor's calculation	Decisions		
Name (F)	No. tablets	EUR/day*	Take into consideration	Dilemma	Tablets	EUR/day
ACEpril ^(F)	2/day	1,60	Guidelines?	Take all of them?	9	5,02
Betanolol ^(F)	2/day	0,81	Side effects?	Just one?	1	~ 0,56
Nonusodipin ^(F)	1/day	0,68	Interactions?	Any combination?	~ 2	~ 1,89
Reservosartan ^(F)	1/day	1,17	Compliance?			
Digiforget ^(F)	1/day	0,09	Budget?			
Aquatonal ^(F)	1/day	0,29				
Spironosubtil ^(F)	1/day	0,38				

(F) = fantasy (no trademark)

* = calculated from the German „Rote Liste 2003“

Von Seiten der Hausärzte werden drei Schwierigkeiten gegenüber leitliniengeförderten Mehrfachbehandlungen angeführt: Die mögliche Zunahme von Wechsel- und Nebenwirkungen, eine nicht begründete ökonomische Mehrbelastung und die mit der Zahl der Tabletten abnehmende Compliance ihrer Patienten, die ebenfalls zur Verschwendung von wichtigen Ressourcen führe. Der Autor fand bei der Mehrzahl seiner Hausbesuche umfangreiche Medikamentenbehälter mit ungenutzten Präparaten – was die Erfahrung vieler Kollegen bestätigt, die sich Gedanken über diese Verschwendung zu Lasten der Gemeinschaft machen.

Das Problem der unerwarteten negativen Interaktion taucht schon bei der Planung von RCTs auf, wenn zu einer evidenzbasiert gesicherten Therapie neue Substanzen hinzukommen sollen. Aus ethischen Gründen wird es meist als unzulässig angesehen, eine gesicherte

² [Nicht im Text:] Trotz wichtiger Kontroversen über die stark unterschiedlichen Nebenwirkungen innerhalb der Klasse der Kalziumantagonisten gibt es bis heute nicht einmal eine vergleichende Studie zwischen den Untergruppen der Dihydropyridine (DPH) und der Nicht-DPH-Präparate (Holzgreve H (2003) Differenzierung und Bewertung der Kalziumantagonisten in der Therapie der arteriellen Hypertonie. Internist 44: 483–490).

und bewährte Therapie zugunsten des Vergleichs mit einem neuen Prinzip in einem Studienarm zu unterlassen. Das gilt fast genauso, wenn das alte und neue Medikament direkt verglichen werden, außer wenn es eine plausible Ungewissheit über die Wirkung beider gibt. Wenn der individuelle Studienarzt von einer Therapie überzeugt ist, darf er sie seinem Patienten nicht wegen einer Randomisierung vorenthalten, er darf also an der RCT nicht teilnehmen. Die einzige ethische Lösung des Problems wäre es, das Konzept der genuinen Ungewissheit auf die der gesamten Forschungsgemeinschaft auszudehnen, als Basis für ein „klinisches Gleichgewicht“ (Freedman 1987)⁽¹⁹⁾ (das natürlich nicht identisch sein darf mit dem Zustand der Ungewissheit zwischen zwei pharmazeutischen Firmen).

Um solche ethischen Probleme zu umgehen, entsteht ein Trend zu Kombinationen, beispielsweise einer neuen mit einer bewährten Substanz gegen das bewährte Medikament allein. Oder es werden mehrere Kombinationsstufen in Seitenarmen von RCT geprüft – woraus wiederum eine Tendenz zur Kombination vieler Medikamente und zu komplexen Therapieempfehlungen resultiert. Werden zwei Substanzen gemeinsam getestet, dann meist in einer studienbedingt relativ starren Dosierung, die später für die Patienten mit besonderem Risiko (Niereninsuffizienz, Diabetes, Lebererkrankungen etc.) so nicht übernommen werden darf – und als dafür ungeprüft bezeichnet werden muss. Bei solchen Patienten ist die Wahrscheinlichkeit für unvorhergesehene Nebeneffekte bei sonst bewährten Behandlungsarten besonders groß (Califf a.a.O.)⁽⁵⁾.

Weiterhin gibt es zu wenige direkt vergleichende oder „noninferiority“ Studien mit dem Ziel herauszufinden, ob ein Medikament dem anderen unterlegen ist (oder nicht), oder ob andere Vorteile bestehen wie geringere Toxizität, leichtere Verabreichung und/oder niedrigere Kosten. Bei 6.095 Patienten mit Myokardinfarkt demonstrierte die ASSENT-3 Studie einen neuen Weg, indem die Patienten in einer offenen, randomisierten Studie mit verschiedenen komplexen therapeutischen „cocktails“ behandelt wurden (The ASSENT-3 Investigators 2001)⁽⁴⁴⁾. In der Kombination mit einer thrombolytischen Therapie wurde unfraktioniertes Heparin gegen niedermolekulares Heparin oder Abciximab getestet mit dem Ziel zu klären, ob die Kombination mit niedermolekularem Heparin (Enoxiparin) ein geeigneteres alternatives Reperfusionssystem zu Abciximab darstellt als unfraktioniertes Heparin – um auf dieser Basis eine „endgültige“ RCT zu planen (was jetzt geschieht). Ein „noninferiority trial“, das sich trotz der Schwierigkeiten bei der Interpretation und dem wahrscheinlichen Mangel an Sponsoren zur Nachahmung empfiehlt (DeMets and Califf 2002)⁽¹¹⁾.

Nebenwirkungen der randomisierten kontrollierten Studien selbst

Zur Zeit des *Cardiac Arrhythmia Suppression Trial* (CAST) glaubten eigentlich alle Kardiologen, dass ventrikuläre Arrhythmien für den plötzlichen Herztod prädiktiv seien und man durch Unterdrückung dieser Arrhythmien dem „Sudden Death“ vorbeugen könne. Doch die erfolgreiche Beseitigung von Arrhythmien erwies sich als falscher Surrogat-Parameter (DeMets and Califf 2002)⁽¹⁰⁾. Die CAST-Studie zeigte ein dramatisches Versagen des Outcome-Parameters: Sehr früh mussten zwei Arme der Studie wegen eines hochsignifikanten Anstiegs plötzlicher Herztodfälle und der allgemeinen Letalität in den Behandlungsgruppen (nicht Placebo!) abgebrochen werden (Ruskin 1989)⁽³⁹⁾.

Alle Kliniker haben daraus gelernt, stets und primär an die klinischen Endpunkte der Lebensverlängerung, der Vermeidung akuter kardialer Ereignisse und der Verbesserung der Lebensqualität zu denken. Andere Endpunkte wie Arrhythmiebekämpfung oder Blutdrucksenkung sind als Surrogate nur unter strikten Vorsichtskriterien bei frühen Phasen oder begrenzter Fragestellung zulässig (DeMets a. a. O.)⁽¹⁰⁾ – aber bleiben immer Surrogate. Hier liegen mit die wichtigsten ethischen Fragestellungen und Aufgaben für die „Institutional Review Boards“ (IRB).

Aber auch bei Studien, die sich auf die „true clinical outcomes“ konzentrieren, sind Studienabbrüche wegen unterschiedlicher Entwicklung in den Studiengruppen nicht selten. P. Poole-Wilson hat kürzlich die „Annual mortality in recent trials in heart failure“ analysiert und Letalitätsunterschiede von 0,4 bis 8,0 Prozent bei Todesfällen in den Gruppen

„Control“ und „Drug“ gezeigt. Bei einem Studienumfang von mehreren Tausend Patienten kann man kalkulieren, was diese Zahlen bedeuten (Poole-Wilson 2003)⁽³²⁾. Auch wenn „Control“ die bisher als Standard angesehene Therapie bedeutet und vorausgesetzt werden darf, dass die Studienabbrüche sofort bei Erreichen statistischer Differenzen erfolgen (DeMets and Califf 2002)⁽¹¹⁾, kann man aus solchen Zahlen Konsequenzen ableiten: Nicht nur sollte man keine Studie ohne Mitwirkung einer Ethikkommission beginnen, sondern sich bereits im Vorfeld der Überlegungen fragen, in welchem Umfang der erzielbare wissenschaftliche oder wirtschaftliche Gewinn die kalkulierte Studiensterblichkeit rechtfertigt.

Von großer ethischer Bedeutung ist der Unterschied zwischen Studien, die neue Therapien testen, und zum Beispiel „noninferiority trials“ mit früher eingeführten Medikamenten: Im Falle einer neuen therapeutischen Maßnahme oder eines neuen Medikaments sollten schon weniger schwerwiegende Ereignisse zum Abbruch der Studie Anlass geben, während eine Neubeurteilung längst eingeführter Behandlungsformen deutlichere Evidenzen benötigt – wie es bei CAST leider der Fall war (DeMets and Califf 2002)⁽¹¹⁾.

Geltungsdauer von Leitlinien. Was meint wissenschaftliche Kreativität?

Im Idealfall sollte der Bedarf der Patienten Anlass für neue Studien sein. Diese wären dann indiziert, wenn neue Therapieformen oder neue Medikamente zu erproben sind, welche die geltenden Leitlinien als überholt erscheinen lassen. Die großen kontrollierten und randomisierten Studien verlangen heute eine Finanzierung, die wohl nur von großen Unternehmen getragen werden kann. Dabei sollte akzeptiert werden, dass die Unternehmen ihre Kosten und Erträge zu kalkulieren haben. Man beobachtet allerdings, dass neue Studien besonders dann gefragt sind, wenn Patente auslaufen und die Einführung von Generika für die forschende Firma einen Preisverfall befürchten lässt. Im Pharmajargon spricht man von einer „Pipeline“, auf der sich ausgereifte neue Substanzen befinden, die bei Bedarf auf die „Studienlinie“ gelangen. Ein aktuelles Beispiel gab die *Neue Zürcher Zeitung (NZZ)* vom 13.02.03: Bei dem bisherigen *Blockbuster* [„Knüller“] *Augmentan*[®] des Pharmakonzerns *GlaxoSmithKline* laufe der Patentschutz aus, jedoch bestehe für die Aktionäre eine „*vielversprechende Produkt-Pipeline*“.

Das sind wirtschaftlich korrekte Abläufe, aber der forschende Arzt muss sich fragen, ob seine wissenschaftlichen Neigungen weitgehend von der Taktgebung der Industrie abhängen sollen. Der Gegensatz wird manifest, sobald unabhängige Institutionen eine Studie zur Behandlung einer Volkskrankheit veröffentlichen, wie kürzlich die durch öffentliche Gelder [NIH] finanzierte ALLHAT-Studie (ALLHAT 2002)⁽¹⁾. Wenn darin ein älteres (und preiswertes) Medikament wie *Chlorthalidon* mit hoher Evidenz wirksam gefunden wird, läuft das gegen die wirtschaftlichen Interessen anderer Hersteller. Eine mit ideologischen Argumenten aufgeladene Diskussion unter dem bezeichnenden Thema „*Evidence-Based Medicine or Medicine that Suggests Evidence?*“ wurde dementsprechend eingeleitet und ist noch nicht abgeschlossen (Messerli 2002)⁽²⁸⁾. Es wäre wünschenswert, wenn dabei auch Gründe für den Mangel an Studien über preiswertere Therapien in der Kardiologie zur Sprache kommen könnten.

Von der EBM unterversorgte Bereiche in der Kardiologie

In wichtigen Teilen der Kardiologie fehlt es an EBM bzw. RCT. Dazu gehören die pädiatrische Kardiologie, mit weltweit weniger als 30 RCTs, die Behandlung von Erkrankungen der Herzklappen, die Pulmonale Hypertonie und die Endokarditis (Califf and DeMets 2002)⁽⁶⁾.

Diese Aussage gilt für große Teile der Diagnostik, früher ein Stolz der Kardiologie. Bei therapeutischen Maßnahmen sind große randomisierte Studien in der Lage, auch kleine Unterschiede nachzuweisen; deshalb findet die Methode der RCT breite Anwendung. Dagegen haben Studien über diagnostische Untersuchungsmethoden sehr häufig das Problem der zu geringen Prävalenz, die beim Hausarzt noch wesentlich kleiner ist als beim Spezialisten (Praetorius 1992)⁽³³⁾; Sox et al. 1990)⁽⁴³⁾, und extrem gering beim „Screenen“ gesunder Popu-

lationen. Dieses Problem kann in ein unlösbar scheinendes Dilemma der Evidenzbasierten Medizin selbst führen, wie kürzlich beim Mammographie-Screening (Goodman 2002)⁽²⁰⁾.

Diagnostische Daten als Grundlage einer RCT beruhen oft auf Methoden, die für die Definition der therapeutischen Ziele und selbst der Pathophysiologie zu grob sind. Ein hoher Evidenzgrad für die therapeutische Maßnahme kann dann täuschen. Beispielsweise sagt für den wichtigen Risikofaktor Fettsucht der Body Mass Index, obgleich Grad B (Yusuf et al. 2002)⁽⁴⁹⁾, wenig über den amount of excess body fat und die ebenso wichtige regionale Fettverteilung aus.

Die Definition der ST-Streckensenkung im Belastungs-EKG bei koronarer Herzkrankheit (KHK) beruht bis heute mehr auf Erfahrung als auf Standards von Kardiologischen Gesellschaften (Fletcher 2001; Trappe and Löllgen 2000)^(18;46). Letztere basieren nicht auf prospektiven Megastudien, sondern auf vielen Veröffentlichungen, die sich über Jahrzehnte der wissenschaftlichen Kardiologie verteilen. Die Wahl eines Grenzwertes³ wird durch die Tatsache kompliziert, dass es für das Belastungs-EKG keine evidenzgestützten Werte gibt, die eindeutig Gesunde von Kranken trennen. Statistisch gesehen ist die Beurteilung von Verlagerungen der ST-Strecke von der Spezifität und Sensitivität des Tests abhängig. Der prädiktive Wert der Untersuchung wird in hohem Maße von der Prävalenz der Erkrankung in der jeweiligen Untersuchungsgruppe beeinflusst (vgl. Bayes' Theorem (Praetorius 1992)⁽³³⁾). Eine Angina pectoris, die bei der Belastungsuntersuchung auftritt, hat eine höhere Treffsicherheit in der Voraussage einer Koronarstenose als das EKG, noch höher freilich wenn zugleich eine ST-Senkung vorliegt - und umgekehrt. In dieselbe bzw. konträre Richtung wirken die Risikofaktoren der KHK sowie Alter und Geschlecht der Patienten. Zusammengefasst handelt es sich bei der Verdachtsdiagnose vor Herzkatheteruntersuchung um eine bunte Mischung mehr oder weniger „weicher“ Faktoren in Kombination mit der klinischen Erfahrung und der jeweiligen beruflichen Situation des Untersuchers (Sox et al. 1990)⁽⁴³⁾.

Ein weiteres Beispiel der Erzeugung klinischer Unsicherheit sind die Leitlinien der DGK zur Interventionellen Koronartherapie (Erbel et al. 1997)⁽¹⁶⁾, welche die therapeutischen Indikationen nach Graden A-C (leider eigenen Kategorien der DGK, einschließlich eines „Konsensus“ bei der A-Kategorie) bewerten und die Stenosegrade von Koronarstenosen mit Hilfe einer Modifikation der AHA/ACC-Kriterien. Allerdings bleibt weltweit die Definition bzw. Quantifizierung von Koronarstenosen bis heute einer subjektiven Beurteilung überlassen, die menschlich verstehbare Tendenzen hat: Bei exemplarischen Audits der deutschen ALKK haben die Untersucher die Stenosen vor der PTCA im Mittel um 4,7% überschätzt und nach der Behandlung um 3,5% unterschätzt (Praetorius 1999)⁽³⁵⁾. Natürlich wurden viele Anstrengungen unternommen, die Objektivität dieser auf Schätzung beruhenden Methoden zu verbessern, leider ohne allgemein akzeptierte Lösungen. Die diagnostische Basis vieler wichtiger therapeutischer RCT wird wohl noch für längere Zeit unterrepräsentiert bleiben.

Systematische Verzerrungen

Die überwältigende Mehrheit internationaler Publikationen und auch von Leitlinien führt schon deshalb zum Ungleichgewicht, weil es mehr möglich ist, das Wichtige vom Unwichtigen zu unterscheiden, und die Kontrolle der methodischen Kriterien des Gelesenen aus dem gleichen Grund höchstens exemplarisch erfolgt. In vielen Fällen hat der Leser den Eindruck, dass wirtschaftliche oder andere Interessen die Ergebnisse einseitig, ja fragwürdig machen – es könnte eine *publication bias* vorliegen. Als gesichert gelten dabei folgende Erkenntnisse: Ein positives Ergebnis einer Studie hat eine höhere Wahrscheinlichkeit publiziert zu werden, sowie in englischer Sprache und mehrfach zur Veröffentlichung zu kommen. Die Gefahr, nicht publiziert zu werden, steigt mit dem Einflussgrad interessierter Sponsoren und der Entfernung der Resultate vom gewünschten Ergebnis. Besser ist die Chance so genannter „high-profile trials“ mit negativen *outcomes* (z. B. CAST, s. oben), in führenden Zeitschriften publiziert zu werden, weil sie für die gesamte Kardiologie von hoher

³ [Nicht im Text:] Die Definition des J-Punktes im EKG und seines Einflusses auf die Beurteilung einer ST-Senkung ist keineswegs präzise (z. B. ab >0,1 oder >0,2 mV 60-80 msec nach dem J-Punkt).

Bedeutung sind und die Herausgeber sie deshalb ohne Rücksicht auf die Ergebnisse der Studie veröffentlichen (DeMets and Califf 2002)⁽¹¹⁾. Das Problem einer statistischen *Bias* bleibt dennoch, wenn „negative“ RCT mit weniger öffentlichem Aufwand geringere Publikationschancen haben.

Zur Vermeidung der Publication Bias wurde in den USA ein Register aller RCT, die „unterwegs“ sind, begründet - um wenigstens ihre Existenz bekannt zu machen. Dieses Register wird an der National Library of Medicine eingerichtet (DeMets a.a.O.)⁽¹¹⁾.

Allgemeine Bemerkungen. Mögliche Konsequenzen.

Zur Evidenz der EBM selbst

Es ist deutlich geworden, dass die Versorgung individueller Patienten in der Kardiologie durch die EBM nicht nachweisbar verbessert werden konnte. Der umfassende Anspruch, wissenschaftlich gesicherte Resultate in objektiver, übersichtlicher und zugleich für die Praxis geeigneter Form zur Verfügung zu stellen, ist noch nicht realisiert. Auch die erhoffte Vereinfachung der Kommunikation unter den Ärzten oder zwischen Arzt und Patient ist nur in Einzelfällen wirksam geworden.

Es scheint eine Spezialität der europäischen Leitlinien-Kommissionen zu sein, dem Benutzer nach differenziert ausgearbeiteten Literaturübersichten und Metaanalysen Empfehlungen anzubieten, denen jedoch eine dezidierte Zusammenfassung mit praktischen Anleitungen fehlt. Es gehört in diesen Zusammenhang, dass auch die klare Kennzeichnung (also Bezifferung) nach internationalen Standards in vielen Empfehlungen fehlt, beispielsweise in Bertrand et al. (2002)⁽⁴⁾ oder Erdmann and Hoppe (2001)⁽¹⁷⁾. Sie werden nicht schon deshalb zu Leitlinien im Sinne des Wortes, wenn sie so genannt werden.

Es gibt noch keine RCT über die Auswirkungen der EBM selbst. Es ist jedoch festzustellen, dass die EBM die praktisch arbeitenden Mediziner statt der erhofften Vereinfachungen mit zusätzlicher Komplexität belastet hat (Vogd 2002)⁽⁴⁷⁾, wie man am Problem der Medikamentenwahl (Tab. 3) sehen konnte.

Interessen durchsichtig machen, neues Vertrauen bilden

Die Ärzte in Klinik und Praxis erhalten fast täglich Empfehlungen und Leitlinien, die von pharmazeutischen Firmen an sie verteilt werden. Das Bewusstsein der Steuerung durch Interessen ist bei den meisten Ärzten vorhanden. Es wird verstärkt, wenn selbst Leitlinien der AHA in die Kritik geraten, wie beispielsweise bei der eigenartigen Höherstufung der Therapie des Schlaganfalls durch Thrombolyse von Iib („optional“) nach Ia durch ein Expertengremium, dessen finanzielle Unabhängigkeit jetzt umstritten ist: Die Sache wurde weniger eigenartig, als eine Spende an die AHA in Höhe von 11 Mill. Dollar durch den betreffenden Pharma-Hersteller bekannt wurde (Lenzer 2002)⁽²⁷⁾, sowie die Tatsache, dass 6 von 9 Mitgliedern der Leitlinienkommission finanzielle Beziehungen zu der Firma haben, die sie nicht deklariert hatten (Smith 2002)⁽⁴²⁾. Die kritisierten Experten argumentieren, dass solche Beziehungen eher eine Belobigung verdienen (!), und sie nennen es eine illusionäre „extreme financial correctness“ (J.L. Saver in ⁽²⁷⁾), wenn die Forderung nach Leitlinienerstellung durch finanziell unabhängige Experten erhoben wird.

1993 veröffentlichte das *International Committee of Medical Journal Editors* eine Handlungsrichtlinie über Interessenkonflikte. Bis heute werden solche Interessenkonflikte in den meisten Zeitschriften jedoch nur selten deklariert, wie etliche Studien gezeigt haben (Smith a.a.O.). In Deutschland bieten einige Zeitschriften eine kurze Erklärungsformel an, in der freilich exakte Angaben zur Rolle des Sponsors fehlen, wie am Beispiel der ERICA (siehe oben)⁽²⁾ gezeigt wurde. Im Jahre 2002 führte das *British Medical Journal* (BMJ) eine randomisierte Studie über den Effekt einer Deklaration von Interessenkonflikten auf die Leser der Zeitschrift durch. Zwei Gruppen von Lesern (n = 170) erhielten dieselbe Publikation, aber in der Gruppe I mit, in Gruppe II ohne die Deklaration spezieller Interessen (die Autoren seien Angestellte eines fiktiven Unternehmens oder sogar Aktionäre). Die Leser der

Gruppe I teilten mit, dass sie den gleichen Text für signifikant weniger interessant, wichtig, bedeutsam, valide und glaubwürdig halten als die Leser der Gruppe II (Choudry et al. 2002)⁽⁷⁾.

Die AHA-Schlaganfallaffäre zeigt, dass auch bei offenkundig unabhängigen Leitlinien Verdachtsmomente auftreten können. Wie oft? Choudry hat diese Frage in einer repräsentativen Umfrage bei 192 Autoren von 44 „Clinical Practical Guidelines“ (CPG) zu klären versucht: 81 % der Autoren je CPG hatten Beziehungen zur jeweils zuständigen pharmazeutischen Industrie. Es scheint zwingend notwendig, die Offenlegung finanzieller und anderer Interaktionen zur Regel zu machen und einen formalen Prozess der Diskussion solcher Konflikte in der Autorengruppe bereits vor der Entwicklung der Leitlinien zu installieren (Choudry et al. 2002)⁽⁸⁾. Es ist nicht prinzipiell moralisch schlecht, solche Interessenkonflikte zu haben, wie sie in der Medizin wahrscheinlich universell sind. Ein Problem entsteht nur, wenn man sie nicht deklariert, weil dadurch das Erscheinungsbild des unabhängigen Wissenschaftlers oder Leitlinienverfassers untergraben wird.

Publikationsaffären haben einen negativen Beitrag zum Zustand der ethischen Standards geleistet, der durch einen dramatischen Niedergang der allgemeinen Moral in der praktischen Medizin charakterisiert ist. Bei einer offiziellen Umfrage unter britischen Allgemeinärzten (general practitioners (GP)) stellten zwei Drittel von ihnen fest, die Moral sei niedrig oder sehr niedrig, und ebenso dass sie niedriger sei als fünf Jahre zuvor (British Medical Association. National survey of GP opinion. London: BMA, 2001; vgl. auch Pendleton and King 2002⁽³⁰⁾). Einer der Gründe für diese Entwicklung dürfte das Versagen der Medizinischen Organisationen und Gesellschaften sein, die es nicht fertig bringen, jene moralischen Grundwerte zu deklarieren, die als handlungsleitende Prinzipien wirken und die ihren Mitgliedern in diesem Sinne Halt geben könnten (Pendleton a.a.O.).

Werden moralischen Werte nicht öffentlich diskutiert, wird sie der einzelne einfach aus dem Verhalten der jeweiligen Gesellschaft erschliessen (Pendleton a.a.O.) – und das sieht nicht immer gut aus, wie z. B. die AHA-Kontroverse. An solchen Ereignissen und aus ihrer eigenen Praxis formen vermutlich viele Kliniker und Praktiker ihre sehr privaten Ethiken, mitgeformt von tiefer Resignation. Wechselnde moralische Prinzipien ebenso wie ständig neue und vergehende evidenzbasierten Leitlinien passen unter dem Signal „anything goes“ zu einer neoliberalen Wirtschaftsgesellschaft der vielen Einzelnen, die dem immer mehr beschleunigten Umsatz pharmazeutischer Produkte dient.⁴ Einige Firmen sind sich der Gefahr bewusst, dass diese Marktpolitik auf die lange Sicht zu einer allgemeinen Deprofessionalisierung führen muss, einschließlich eines nachfolgenden Niedergangs der Forschung und in letzter Konsequenz der industriellen Produktion selbst. Die Vision einer Neuformierung um firmenspezifische Wertbegriffe wird gefordert (*moral leadership*). Aber bevor man eine Vision bekommt, ist es besser Transparenz zu praktizieren, damit alle erkennen, wofür die Gesellschaft steht. (siehe Pendleton a.a.O.).

Transparenz von Leitlinien-Komitees

Die strengen Gesetze der RCT sind nicht mit den Regeln der Zusammenstellung von unabhängigen Leitlinien-Komitees identisch (vgl. Lenzer)⁽²⁷⁾. Selbstverständlich müssen Experten, die Mitglieder eines solchen Komitees werden, ihre oft komplizierten Beziehungen zu kommerziellen Anbietern deklarieren (Spenden für die Forschung, Vortragseinnahmen, Beratungsgelder, wirtschaftliche Teilhabe etc.). Die Untersuchung von Choudhry nennt drei Punkte: Angemessene Deklaration aller finanziellen Interessenkonflikte (1.), ein formalisierter interner Prozess zu dieser Frage (2.) und (3.) die komplette Offenlegung aller auch nur möglichen Interessenkonflikte für die Leser der Leitlinien (Choudhry 2002)⁽⁸⁾.

Ein weiterer formaler Prozess dient dem Erreichen eines korrekt entwickelten Konsensus der Experten im Leitlinienkomitee. Darin wird ebenso der Vorgang der Wertung vorliegender Studien, Übersichtsartikel etc. definiert wie der Weg zur eigentlichen Entscheidung, der neben formalen Beschlussregeln auch die dynamischen Prozesse einer solchen Experten-

⁴ Liberalistische Flexibilität der vielen Einzelnen. Sie fügt sich sehr gut in die globalisierende Ökonomie ein. Der Umsatz neuer Methoden wird größer, aber durch einen Verlust an Verantwortung erkaufte. [nicht im engl. Text].

gruppe selbst kontrollieren soll. In Deutschland wird ein derartiges Schema von der AWMF (*Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften*) angeboten. Die bisher unbefriedigende oder teilweise sogar fehlende Evidenzklassifikation in deutschen und europäischen kardiologischen Leitlinien (vgl. Lauterbach 2002)⁽²⁶⁾ könnte ihre Ursache in der europäischen Praxis informeller und unkontrollierter Expertengruppen haben, in denen niemand bereit ist, unter Verzicht auf eigene „hausgemachte“ wissenschaftliche Theorien oder Forschungspraktiken zu kooperieren.

Regeln für die Durchführung von RCT und die Entwicklung von Leitlinien, in denen Interessenkonflikte, der finanziellen Ausgleich für klinische Forschungsarbeiten und die ethischen Verpflichtungen bei Experimenten mit Menschen formuliert sind, werden schon seit langem gefordert (Topol et al. 1997)⁽⁴⁵⁾. Jeder, der in einer RCT Patienten behandelt, muss beim Angebot von Zahlungen über den Bereich von Kostenerstattung für effektive Leistungen hinaus zurückhaltend werden, ebenso wenn dies für die Aufnahme bestimmter Patienten in die Studie angeboten wird; analoges gilt für Leitlinienkomitees (DeMets a.a.O. ⁽¹¹⁾). Eine detaillierte Übersicht zu diesen Fragen findet sich bei der „Task Force on Financial Conflicts of Interest“ des AAMC (2001)⁽³⁾.

In Deutschland wurden von verschiedenen Regierungen über 14 „Gesundheitsreformen“ versucht. Alle scheiterten, weil bis heute zu viele Interessengruppen und Parteien sich an diesem Spiel beteiligen. Es scheint nur noch ein Traum zu sein, dass eine wirklich abschließende Gesundheitsreform noch einmal kommt (Praetorius 2003)⁽³⁶⁾. Einer der Traum-inhalte wäre eine faire Interessenbalance zwischen Patienten und ihren Hausärzten einerseits und den Spezialisten und der Gesundheitsindustrie auf der anderen Seite – zur Bildung neuen Vertrauens und Implementierung des Geltungsanspruches verbindlicher moralischer Normen, ohne zugleich den Fortschritt der Forschung zu behindern. Wenn dieser Traum nicht erfüllt wird (und die Zeichen stehen ungünstig), reagieren Politiker gerne mit der Alternative strikter Regulierung der Praxis. Es ist eine Horrorvorstellung: Die tödliche Umarmung und Strangulierung der Medizin durch eine bürokratische EBM, wie sie bereits auf dem Feld der „diagnose related groups“ (DRG) stattfindet.

EBM als neues Paradigma der Medizin?

EBM versteht sich einerseits als Zukunftsprojekt einer verbesserten Wissenschaftlichkeit und insofern als Modernisierungsbewegung. Andererseits will sie die wissenschaftliche Basis praktischen ärztlichen Handelns bilden. Insgesamt ist EBM ihrer Natur nach eher konservativ und affirmativ und für biomedizinische Innovationen primär nicht kompetent (Raspe 2001; Rogler and Schölmerich 2000)^(37,38). Insofern ist die gelegentlich zu hörende Behauptung, dass EBM das neue Paradigma der medizinischen Wissenschaft sei, logisch verfrüht. Ein Paradigma im Sinne von Thomas Kuhn (Kuhn 1962)⁽²⁴⁾ entsteht neu im Zusammenhang mit wissenschaftlichen Umwälzungen (oder geht als veraltete Tradition zugrunde); man kann es bekanntlich nicht herbeiwünschen, sondern allenfalls über seine gegenwärtige Geltung oder über seinen Niedergang diskutieren. Die Diskussion über ein zukünftiges Paradigma gehört demnach eigentlich in das Gebiet der *Science Fiction*. Das Reden darüber im wissenschaftlichen Kontext hat einen performativen Beiklang, der zur Verschleierung der Kritik realer Probleme beiträgt. Jede Form von *Wissensinszenierung* ist für die EBM eine Gefahr, weil sie zu einer „Ideologie der Rationalität“ und damit von der Realität weg führen kann, wie kürzlich Werner Vogd gezeigt hat (Vogd 2002)⁽⁴⁷⁾.

Die in Leitlinien formulierten Standards verstärken die Anforderungen an den einzelnen Arzt und addieren sich zu den Verschlüsselungsarbeiten in Klinik und Praxis, die ihm bereits zugemutet wurden. Er könnte auf die Herausforderungen mit einer Stärkung seiner Autonomie antworten. Leider wird es jedoch im Gefolge einer politisch überregulierten EBM mit größerer Wahrscheinlichkeit zu einem technokratischen Rückzug kommen, zur *Abstraktion vom konkreten Patienten* und zum *Verlust an ärztlicher Autonomie und Professionalität* (Vogd a.a.O.), mit progredientem Verantwortungsverlust („*medical deresponsibilization*“ nach Sureau 1995, siehe ⁽³⁴⁾). Aber wir sollten einräumen, dass EBM auch die Chance enthält, die starken kommerziellen

Interessen in der Medizin zu kontrollieren und zugleich zu einer Verbesserung der ärztlichen Fortbildung und Praxis zu führen. Sie ist eine Methode, die wie andere Methoden von unserer individuellen Ehrlichkeit, Unabhängigkeit und wissenschaftlichen Einstellung abhängt.

References

1. ALLHAT Collaborative Research Group (2002) Major Outcomes in High-Risk Hypertensive Patients Randomized to Angiotensin-Converting Enzyme Inhibitor or Calcium Channel Blocker vs Diuretic. *JAMA* 288: 2981-2997
2. Antoni DH, Beuckelmann D, Hambrecht R, et al. (2002) Therapie mit Beta-Adrenorezeptorantagonisten bei der chronischen Herzinsuffizienz. *Deutsche Medizinische Wochenschrift* 127: 2459-63
3. AAMC Task Force on Financial Conflicts of Interest in Clinical Research (2001): Protecting Subjects, Preserving Trust, Promoting Progress. Policy and Guidelines for the Oversight of Individual Financial Interests in Human Subjects Research. <http://www.aamc.org/members/coitf/firstreport.pdf>
4. Bertrand ME, Simoons ML, Fox KA, et al. (2002) Management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. *European Heart Journal* 23: 1809-1840
5. Califf RM, DeMets DL (2002) Principles From clinical Trials Relevant to Clinical Practice: Part I. *Circulation* 106: 1015-1021
6. Califf RM, DeMets DL (2002) Part II. *Circulation* 106: 1172-1175
7. Chaudhry S, Schroter S, Smith R, Morris J (2002) Does declaration of competing interests affect readers' perceptions? A randomised trial. *BMJ* 325: 1391-2
8. Choudhry NK, Stelfox HT, Detsky AS (2002) Relationships Between Authors of Clinical Practice Guidelines and the Pharmaceutical Industry. *JAMA* 287: 612-617.
9. Davidoff F, et al. (2001) Sponsorship, Authorship, and Accountability. *New England Journal of Medicine* 345: 825-827
10. DeMets DL, Califf RM (2002) Lessons Learned From Recent Cardiovascular Clinical Trials: Part I. *Circulation* 106: 746-751
11. DeMets DL, Califf RM (2002) Part II. *Circulation* 106: 880-886
12. Deng MC et al., COCPIT Study Group (2000) Effect of receiving a heart transplant. *BMJ* 321:540-5
13. Deutsche Forschungsgemeinschaft (1999) Klinische Forschung – Denkschrift. ©2002-3 <http://www.dfg.de>
14. Dierks C (2003) Legal implications of guidelines. *Deutsche Medizinische Wochenschr* 128:815-819
15. Dissmann W, Ridder M de (2002) The soft science of German cardiology. *Lancet* 359: 2027-29. Discussion: *Lancet* 360: 1694-96.
16. Erbel R, Engel HJ, Kübler W, et al. (1997) Interventionelle Koronartherapie. *Zeitschrift für Kardiologie* 86: 1040-1063.
17. Erdmann E, Hoppe UC (2001) Leitlinien zur Therapie der chronischen Herzinsuffizienz. *Zeitschrift für Kardiologie* 90: 218-237
18. Fletcher GF (2001) Exercise Standards for Testing and Training. *Circulation* 104: 1694-1740
19. Freedman B (1987) Equipoise and the ethics of clinical research. *New England Journal of Medicine* 317: 141-5
20. Goodman SN (2002) The mammography dilemma: A crisis for Evidence-Based Medicine? *Annals of Internal Medicine* 137:363-65
21. Haynes RB, Devereaux PJ, Guyatt GH (2002) Physicians' and patients' choices in evidence based Practice. Evidence does not make decisions, people do. *BMJ* 324: 1350
22. Joost H-G (2003) Welche Interaktionen gibt es bei gleichzeitiger Gabe von ASS und ACE-Hemmern? *Deutsche Medizinische Wochenschrift* 128: 398
23. Kuhlmann E (1999) Between informed consent and denying information: patient information un-der economical constraints. *Ethik Med* 11: 146-61
24. Kuhn TS (1962) Die Struktur wissenschaftlicher Revolutionen. Suhrkamp, Frankfurt am Main.
25. Kunz R, Ollenschläger G, et al. (2000) Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis. Deutscher Ärzteverlag Köln.
26. Lauterbach KW, Redaelli M (2002) Evidence based medicine in der Kardiologie. *Cardio News* 4: 41-42
27. Lenzer J (2002) Alteplase for stroke: money and optimistic claims buttress the "brain attack" campaign. *BMJ* 324:723-726. Saver JL, Kidwell CS, Starkman S: Commentary: Thrombolysis in stroke: it works! *BMJ* 324:727-729
28. Messerli F (2002) Life protection: Learnings from new clinical trials in hypertension. The world after ALLHAT. *Cardiology Update* 2003, Davos.
29. Pearson TA, Mensah GA, Wayne Alexander R, et al. (2003) Markers of Inflammation and Cardio-vascular Disease. *Circulation* 107:499-511. <http://circ.ahajournals.org/cgi/reprint/107/3/499.pdf>
30. Pendleton D, King J (2002) Values and leadership. *BMJ* 325:1352-5
31. Perleth M, Raspe H (1998) Evidenzbasierte Medizin: Möglichkeiten und Grenzen. *Münchener Medizinische Wochenschrift* 140:99-102
32. Poole-Wilson P: Priorities in health care and ethical considerations. Davos Lecture, Febr. 14, 2003
33. Praetorius F (1992) Bayes City: The Objectivity of Indications (Bayes-Stadt ... Ein Plädoyer für den diagnoseführenden Arzt). *Deutsches Ärzte-blatt* 89, 2113-2120
34. Praetorius F, Sahn S (2001) The Ethical Aspect of Regularisation in Medicine. *Ethik Med* 13: 221-242

- 36.35. Praetorius F: Coronary Angioplasty (PTCA) Registry of the German Community Hospitals: Ethical Consequences of an Internal Quality Control. (Selbstbegrenzung als Modell?) *Ethik Med* (1999) 11:89-102
- 37.36. Praetorius F (2003) Nach der finalen Reform - Traum eines Kardiologen. *Deutsches Ärzteblatt* 100: 3212
- 38.37. Raspe H (2001) „Die Heilkunde wird eine Wissenschaft sein, oder sie wird nicht sein“. *Zeitschrift für ärztliche Fortbildung und Qualitätssicherung* 95: 495-501
- 39.38. Rogler G, Schölmerich J (2000) „Evidence-Biased Medicine“ – oder: Die trügerische Sicherheit der Evidenz.. *Deutsche Medizinische Wochenschrift* 125: 1122-1128
- 40.39. Ruskin JN (1989) The Cardiac Arrhythmia Suppression Trial (CAST). *New England Journal of Medicine* 321: 386-388
- 23.40. Schulman KA, Seils DM et al. (2002) A National Survey of Provisions in Clinical-Trial Agreements between Medical Schools and Industry Sponsors. *New England Journal of Medicine* 347:1335-1341
41. Senges J, Gitt AK (2002) Reicht die zur Verfügung stehende Datenmenge und Datenqualität aus, um medizinische Entscheidungen abzusichern? Evidence-based medicine. 68. Jahrestagung der Deutschen Gesellschaft für Kardiologie.
42. Smith R (2002) Making progress with competing interests. *BMJ* 325: 1375-1376.
43. Sox HC, Hickam DH, Marton KL, et al. (1990): Using the Patient's History to Estimate the Probability of Coronary Artery Disease: A Comparison of Primary Care and Referral Practices. *American Journal of Medicine* 89: 7-14.
44. The ASSENT-3 Investigators (201) Efficacy and safety of tenecteplase in combination with enoxaparin, abciximab, or unfractionated heparin. *Lancet* 358: 605-13
45. Topol EJ, Califf RM, Van de Werf F, et al. (1997) Perspectives on large-scale cardiovascular clinical trials for the new millennium. *Circulation* 95:1072-1082
46. Trappe HJ, Löllgen H (2000) Leitlinien zur Ergometrie. *Zeitschrift für Kardiologie* 89:821-37
47. Vogd W (2002) An Increase in Professionalization or Erosion of Medical Autonomy? The Implications of Evidence Based Medicine and the New Functional Elites in Medicine Seen from the Perspective of Systems and Interaction Theory. (Professionalisierungsschub oder Auflösung ärztlicher Autonomie?). *Zeitschrift für Soziologie* 31: 294-315.
48. Von Schacky C (2002) Koronare Herzkrankheit. *Deutsche Medizinische Wochenschr* 127:2429-31
49. Yusuf S, Cairns JA, et al. (2002) Evidence-based Cardiology. *BMJ Books*, London.