

Karlheinz Bayer
Allgemeinarzt und Chirotherapeut

Überlegungen zur evidenzbasierten Medizin (EbM) im Hinblick auf ihre Eignung als Werkzeug, um die bisherige Medizin zu reformieren

„Evidenzbasierte Medizin“ ist zunächst nur ein Begriff, allerdings einer, der schon durch seinen Namen hohe Erwartungen weckt, seitdem eine Publikation aus 1996 die von Sackett begründete Idee in die medizinische Öffentlichkeit gebracht hat¹). Die evidenzbasierte Medizin (EbM) war von Beginn an umstritten und wird sowohl in der Fachpresse²)³) wie in der Laienpresse⁴) kontrovers diskutiert.

Die Zielsetzung der EbM hat sich gewandelt. War EbM ursprünglich eine Technik zur Aufdeckung medikamentöser Risiken und Nebenwirkungen, wird inzwischen mehr und mehr versucht, sie als Instrument zu benutzen, mit dem die Medizin schlechthin reformiert werden soll⁵).

Was steckt hinter dem Begriff EbM?

Die zusammengesetzten Begriffe „evidenzbasiert“ oder „evidence-based“ sind im Deutschen wie im Englischen nicht exakt definierte Kunstwörter. Der Begriff "**Evidenz-basierte-Medizin**" scheint selbsterklärend zu sein. „Evident“ besagt, daß etwas „augenscheinlich“ ist. Durch das Attribut "basiert" wird das Vorhandensein eines Fundaments suggeriert. Es handelt sich jedoch um eine Übersetzung des amerikanischen Ausdrucks **evidence-based medicine**, und im Amerikanischen heißt „augenscheinlich“ „obvious“, während das Wort „evidence“ mit „(offenliegender) Beweis“ übersetzt werden muß. Es ist fraglich, ob es eine bewiesene Medizin, oder eine Therapie auf der Basis unumstößlicher Wahrheiten überhaupt geben kann.

Selbst der Begriff "**Medizin**" ist im Grunde undefiniert.

Es gibt neben der evidenzbasierten auch eine Schul- und eine alternative Medizin und etliche weitere. Das Wort "Medizin" wird sowohl für den gesamten "Arbeitsbereich Medizin" wie für die "konkret verabreichte Medizin" benutzt.

Das "Evidenzbasiertsein" wird rechtfertigt dadurch, daß die EbM sich auf **wissenschaftlichen Studien** gründet. Zudem hat die EbM alle theoretisch verwendbaren Studien in eine Rangfolge gebracht, denen sie dann eine unterschiedliche Qualitätsstufe zuschreibt. Nach welchen wissenschaftlichen Erkenntnissen die Väter der EbM diese Reihenfolge festgelegt haben, ist offen.

Der Aufbau dieser Qualitätshierarchie folgt einem methodischen Regelwerk, das sich **Good Clinical Practice (GCP)** nennt. „Gute klinische Praxis“ klingt im Deutschen, als handele es sich um eine „gute Medizin“, „gut“ scheint dabei wie eine Wertung. Das englische Wort „good“ wird mehr im Sinn von „dem Zweck gut angepaßt“ oder „handwerklich-verfahrenstechnisch gut“ benutzt, und auch „practice“ impliziert, anders als im Deutschen nicht „Praxis“ bzw. „Praktik“, sondern ist ein Ausdruck für die verwendete Methode, die bessere Übersetzung für "practice" heißt „Anwendung“. Der GCP bedient sich anderer „guter Praktiken“, nämlich der Good Manufacturing und der Good Laboratory Practice (GMP und GLP). Schon dieses komplexere Regelwerk zeigt, daß GCP die Richtlinien der Herstellung und Verabreichung von Medikamenten und Medizinprodukten beschreibt.

EbM hat eine Geschichte.

Die Gesellschaften und Firmen, die hinter dem Regelwerk stehen und es pflegen, bestätigen diese Ausrichtung⁶). Historisch sind alle „good practices“ entstanden in der Folge des **Thalidomid-Skandals**, bei dem es durch die Einnahme des Medikaments **Contergan**[®] zu schweren Mißbildungen bei Neugeborenen kam. GCP, GMP und GLP sollten unter dem Eindruck des Contergan-Prozesses gegen die Firma Chemie Grünenthal von 1967-70, zu einer größeren Medikamentensicherheit und zu einer Imageverbesserung der Pharmaproduktion führen.

Stand bei der Beweisführung der klinischen Erprobungsstudien noch die Erkennung auch seltener Nebenwirkungen eines Medikaments im Vordergrund, ist es heute eher so, daß sie viel mehr der Herausarbeitung selbst kleinster Nutzeffekte derselben Medikamente zu dienen hat. Zweifelsohne stellt jede Verbesserung der Medikamentensicherheit und jede Nutzbarkeit auch kleinster positiver Effekte eine Verbesserung der medizinischen Qualität dar, sie ist aber nicht eins zu eins gleichsetzbar mit medizinischer Qualitätsverbesserung schlechthin. Faktisch hat dadurch beim Risikobegriff ein Paradigmenwechsel eingesetzt. Während es in der frühen Phase zwingend erschien, einen Satz zu bilden wie den: „**zu Risiken und Nebenwirkungen fragen Sie Ihren Arzt oder Apotheker**“, sieht der Ansatz heute eher so aus, die Risiken auf dem Feld der Krankheitsbedrohung zu definieren, und selbst für kleinste relative Risikominderungschancen die Empfehlung auszugeben, Medikamente einzusetzen. Dieser Paradigmenwechsel hat erstmals die Studienergebnisse auch einsetzbar gemacht für Werbezwecke. Während man eine kleine Zahl von Nebenwirkungen kaum plakativ bewerben kann, ist dies bei Minderungen auch kleinster Gesundheitsrisiken sehr wohl der Fall. Auch die Vehemenz, mit der die EbM eingeführt wird und die Verwendung der genannten Begriffe, erinnert in weiten Teilen an **werbepsychologische Strategien**. Bei jeder Art von Werbung ist für den Verbraucher von Bedeutung, **wer** wirbt, **was** beworben wird und **mit welchen Argumenten** und zu **welchem Zweck** geworben wird. Verbraucher und Zielgruppe der EbM sind die Therapeuten ebenso wie die Patienten. Die Argumente verdienen eine genauere Betrachtung.

Die evidenzbasierte Medizin wird sehr intensiv von Sozialpolitikern beworben. Begründet wird der **Bedarf an EbM** - offen oder unterschwellig - damit, daß die bisherige Medizin insuffizient sei und deshalb reformbedürftig. Kienle ²⁾ und Jachertz ³⁾ haben im Deutschen Ärzteblatt unter anderem die Befürchtung herausgearbeitet, daß die Methodik der EbM geeignet ist, zwischen der sogenannten Schulmedizin und den alternativen Therapieverfahren zu Ungunsten der Alternativen Behandlungsmethoden angewandt zu werden.

Namentlich im Fall der **Homöopathie** wurde durch EbM-orientierte Studien versucht nachzuweisen, daß die von Ärzten und Patienten beschriebene Wirksamkeit allein das Ergebnis einer subjektiven Sichtweise sei. Das Studiendesign war so ausgelegt, daß nach Beweisen gesucht wurde, daß Placeboeffekte und Spontanheilungen therapeutisch überproportional wirksamer sind als die Therapie selbst. Die Größenordnung des Heilerfolgs liegt in etwa auf dem niedrigen Niveau, wie man es bei der Chemotherapie bei vielen Krebsarten findet.

Im Fall der **Chemotherapie** geht die EbM allerdings den umgekehrten Weg, neben der Spontanheilung und dem suggestiven Therapieerfolg wenigstens Reste einer Verbesserung nachweisen zu können.

Im Fall der Herzinfarkttrisikosenkung durch **Lipidsenker** wurde in einigen Studien die Fragestellung sogar so sehr eingengt auf die Verwendung von Statinen, daß andere Strategien der Prävention wie Rauch-Verzicht, Gewichtsreduktion, mehr Bewegung und weniger Streß überhaupt nicht zu einem statistischen Vergleich des Therapieerfolgs herangezogen werden, obwohl sie um Zehnerpotenzen wirksamer sind ⁷⁾.

Allein die Fragestellung einer Studie entscheidet darüber, ob im einen Fall die Spontanheilung und der Placeboeffekt oder im anderen Fall ein Rest-Heilerfolg sich als zentrale Aussage im Studienergebnis wiederfinden lassen. Ein pragmatischer Ansatz, der primär oder ausschließlich den meßbaren Therapieerfolg im Auge hat, egal ob dieser durch die Methode selbst oder durch einen Placeboeffekt oder durch andere (soziale, soziogene, psychogene, suggestive, zufällige, magische usw.) Phänomene oder Ansätze erzielt wird, kann in einer EbM nicht Fuß fassen. Die EbM erweist sich allein schon durch das Nicht-in-Betrachtziehen aller anderen Therapieformen, als Förderer medikamentenorientierter Therapien. Hierbei ist auch die Qualitätshierarchie der Studien von großer Bedeutung, nach der die prospektive Doppelblindstudie eine höhere „Evidenzklasse“ besitzen soll, als eine Expertenmeinung.

Was macht eine gute Studie aus?

Die Protagonisten der EbM begründen ihre Annahme, EbM führe zu einer besseren Medizin damit, daß alle Therapiemethoden, die Eingang finden wollen, durch Studien abgesichert sein müssen. Es wird die Hypothese aufgestellt, daß sich über den Weg von Studien **Leitlinien** formulieren lassen, welche nach und nach die bestehende Medizin ersetzen sollen. Die Studien wiederum werden von der EbM **Evidenzklassen** unterworfen, quasi Leitlinien für das Design von Studien.

Stufe Ia: Wenigstens eine Metaanalyse auf der Basis methodisch hochwertiger randomisierter und kontrollierter Studien (RCT)
Stufe Ib: wenigstens ein ausreichend großer, methodisch hochwertiger RCT
Stufe IIa: wenigstens eine hochwertige Studie ohne Randomisierung
Stufe IIb: wenigstens eine hochwertige Studie eines anderen Typs, quasi-experimenteller Studie
Stufe III: mehr als eine methodisch hochwertige nichtexperimentelle Studie
Stufe IV: Meinungen und Überzeugungen von angesehenen Autoritäten (aus klinischer Erfahrung); Expertenkommissionen; beschreibende Studien
Stufe V: Fallserie oder eine oder mehrere Expertenmeinungen

Tabelle: Evidenzklassen

Es gibt jedoch keine vorgelegten Beweise, daß gute Studien auch zu einer besseren Medizin führen, auch wenn dieser Denkansatz einleuchtend scheint. Der einfachste und historisch oft gegangene wissenschaftliche Beweis einer These, geht über die Formulierung einer Gegenthese, die zu widerlegen ist. Die Gegenthese zum Regelwerk der EbM heißt, es ist nicht möglich, über den Weg der vorgeschlagenen Studienhierarchie eine bessere Medizin zu erreichen. Hierzu reicht bereits eine eingehende Betrachtung dessen, was eine gute Studie nach EbM-Definition ist.

Als höchstwertige Art der Studie definiert die EbM die **prospektive randomisierte und kontrollierte Doppelblindstudie** (randomized controlled trial = **RCT**). Als niedrigste Evidenzklasse hat die EbM die **persönliche Erfahrung** eingestuft. Diese Kategorisierung geht auf **Cochrane**⁸⁾ zurück, der sich damit Gedanken machte über die **Effizienz der Randomisierung**. Cochranes Überlegungen haben sich ursprünglich nicht mit dem **Stellenwert einer Studie** oder gar mit der Frage beschäftigt, ob die Meinung von Experten einen geringeren Wert hätte als eine randomisierte Studie, und es gibt weltweit keine Studie, die eine Antwort gibt, warum diese und keine andere Hierarchie richtig sein soll⁹⁾.

Den RCT-Studien wurde im System der EbM das Attribut eingeräumt, „**Goldstandard**“¹⁰⁾ zu sein. „Goldstandard“ ist eine Metapher, in der wir bleiben können, wenn im Folgenden die RCT auf ihren „Goldgehalt“ geprüft wird.

Randomisierung (etwas dem Zufall überlassen) scheint ein Ausweg aus der Gefahr, daß Studien immer manipulierbar sind. Mit demokratisch empfundener Freude schreibt Frau Hackenbroch im SPIEGEL 39/2005⁴⁾, daß die EbM dank der multizentrischen, prospektiven und kontrollierten Doppelblindstudien mehr Demokratie in die Medizin bringe, und die Macht und den Einfluß der Chefärzte ("Eminenz-basierte Medizin") einschränke. Es wird die Annahme unterstellt, Demokratie stehe der Wissenschaft und Chefärzte der Manipulation näher. Diese Meinung über die EbM ist weit verbreitet.

Die **Randomisierung** dient in erster Linie dazu, zu Versuchsgruppen zu kommen, die ein möglichst hohes Maß an Vergleichbarkeit besitzen, an denen man im Weiteren unterschiedliche Therapien anwenden kann. Dahinter verbirgt sich einerseits der Wunsch, gleiche Ausgangsbedingungen zu bekommen, so daß die Ergebnisunterschiede möglichst nur den Unterschieden der Therapien zuzuschreiben sind und im Idealfall nicht durch die Vorstellungen des Untersuchers beeinflusst werden.

Weil aber die Randomisierung möglichst gleich geformte Vergleichsgruppen schaffen soll, müssen im Vorfeld Patienten ausgeschlossen werden, insbesondere solche die zu stark von der Norm abweichen.

Der Studiendesigner greift manipulierend ein, definiert Problemfälle, die herausgenommen werden, weil sie das Studiendesign stören würden. Die Gefahr der Künstlichkeit haftet aus diesem Grund allen randomisierten Studien an ¹¹⁾ ¹²⁾. Der Begriff „**Randomisierung**“ täuscht eine Exaktheit vor, die in Wahrheit umso weniger gegeben ist, je intensiver der Versuch unternommen wird, die Randomisierung zu vervollständigen.

Durch die Verwendung der **doppelt blinden Studien** tauchen andere methodische Probleme auf, die vor allen Dingen die **nicht-medikamentösen Therapieformen** betreffen. Während es mehr oder weniger unproblematisch gelingt, gleich aussehende Medikamentendarreichungen zu produzieren, mal mit und mal ohne Wirkstoff, so daß tatsächlich weder der Patient (Blindstudie) noch Patient und Arzt (Doppeltblindstudie) wissen, ob im vorliegenden Fall **ein Wirkstoff oder ein Placebo** ¹³⁾ gewirkt hat, ist dies bei den nicht-medikamentösen Therapien unmöglich. Wie sollte ein Vergleich zwischen medikamentösen und nicht-medikamentösen Therapieformen, etwa der Vergleich zwischen Patienten die Diätmaßnahmen bekommen und solchen, die alternativ Tabletten bekommen, sich im Blindversuch realisieren lassen? „Doppeltblind“ ist zur Bewertung einer nicht-medikamentösen Therapien auch deshalb ungeeignet, weil der Therapeut, als Akupunkteur, Chirotherapeut oder Psychotherapeut, selbst einen Teil der verabreichten Medizin darstellt, der sich aber sicherlich nicht durch ein Placebo ersetzen ließe.

Es stellt sich grundsätzlich die Frage, ob ein Reagenzglasdenken der therapeutischen Wirklichkeit entsprechen kann, und ob eine wirksame Therapie ohne Placebo für möglich und machbar (oder für erstrebenswert) gehalten wird. Das EbM-System bietet zwar formal-methodische Gründe an, warum der Placeboeffekt ausgeschlossen werden soll, bleibt aber eine Erklärung schuldig, warum der durch Placebo herbeigeführte Teil des Heilerfolgs nicht ebenso wertvoll sein soll wie jeder denkbare andere.

„**Prospektiv**“ ist eine Umschreibung dafür, das Studienziel möglichst exakt vorzugeben. Am nützlichsten kann prospektives Vorgehen dann sein, wenn die Subjektivität in der Bewertung des Therapieerfolgs in sehr begrenzten Einzelfragen minimiert werden soll.

Ohne Zweifel gibt eine retrospektive Betrachtung viel Raum für Interpretationen, wodurch die Expertenmeinung überproportional zum Zuge kommt, und damit die Gefahr der Manipulation. Während allerdings eine retrospektive Studie Raum für Manipulationen bieten KANN, erfordert die prospektive Studie, daß bereits vor Beginn gesteuert und interpretiert werden MUSS. Die Qualität der Studie ist deshalb genauso vom Einfluß der Experten abhängig, der einzige Unterschied besteht darin, daß dieser Einfluß im Design versteckt und nicht in der Analyse offensichtlich ist. Die prospektive Studie ist durch die eingengegte Zielvorgabe sogar unflexibler als die retrospektive Analyse, den die ermöglicht im positiven Sinn auch konkurrierende und alternative Interpretationen. Überraschenderweise ist die retrospektive Analyse deshalb sogar demokratischer als jede prospektive Studie.

Diskussion.

Wir befinden uns bereits auf einem Weg hin zu mehr Leitlinien und zu einem System von Therapieempfehlungen. Es ist jetzt notwendig, geeignete Werkzeuge zu entwickeln, diese Leitlinien zu schaffen. Es müssen Studienkonzepte erarbeitet werden, die gleichrangige Bedingungen schaffen für die sprechende und die nicht-medikamentösen Therapien. Wenn allein die EbM als Maßstab herangenommen wird, werden ohne Zweifel in absehbarer Zeit alle nicht-medikamentösen Therapiemethoden den Rang von weniger wertvollen Therapien bekommen. Es ist nicht so, daß die EbM dem selbst gegebenen Anspruch auf größtmögliche Wissenschaftlichkeit gerecht wird. Das goldene Kalb der randomisierten Doppelblindstudien darf nicht mehr unkritisch angebetet werden.

Der Risikobegriff hat sich unter dem Einfluß der EbM auf das individuelle Ereignisrisiko für Krankheiten fokussiert und ist abgerückt vom Risiko, das der Einnahme von Medikamenten anhaftet. Dadurch ist ein gänzlich neuer Präventionsbegriff entstanden, der mit den ursprünglichen und eigentlichen Zielen der Prävention, nämlich eine Therapie unnötig werden zu lassen, total gebrochen hat.

Als ein extrem konsequentes Beispiel der EbM darf die von Wald und Law ins Spiel gebrachte Diskussion über die Polypill¹⁴⁾ angesehen werden. In dieser Studie wird die Überlegung angestellt, ob es durch die konsequente (100%ige) Einnahme einer Pillenmischung möglich sei, die Risiken für Krankheiten auf ein statistisches Minimum zu senken. Wald und Law gehen so weit, daß sie die individuelle Beratung und Therapie im Fall der kategorischen Einnahme ihrer Polypill für überflüssig halten.

Jede Reform der Medizin sollte sich Gedanken machen nicht nur über Rahmenbedingungen, wie das Design von Studien und die gegebenen Therapieempfehlungen, sondern auch über echte Inhalte. Dazu gehört auch der Aspekt der Therapiefreiheit, der in keinem Ansatz der EbM formuliert ist, und es muß darum gehen, was der Patientenwille ist. Ähnlich wie bei der Patientenverfügung, muß auch hier eine Diskussion einsetzen. Die eher demagogische Haltung der EbM und die autoritäre Konsequenz, Leitlinien zu formulieren, sind geeignet, diesen Anspruch eher zu behindern.

Bad Peterstal, Oktober 2005
Dr.Karlheinz Bayer

改善

das japanische Zeichen „Kai-Zen“ -
Inbegriff des Wunschs, fehlerfrei arbeiten zu wollen

¹⁾ Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. **Evidence based medicine: what it is and what it isn't.** BMJ. 1996;312:71-72

²⁾ Kienle,G. **Evidenzbasierte Medizin - Konkurs der ärztlichen Urteilskraft** Dtsch Ärztebl 2003; 100:A 2142-2146

³⁾ Jachertz, Norbert **Bewertung von Therapien: „Korridor der Vernunft“** Dtsch Ärztebl 2005; 102, A-268 / B-217 / C-204

⁴⁾ Hackenbroch,V. **Wir lernen uns zu wehren,** DER SPIEGEL 39/2005 (26.9.2005) Seite 142ff.

⁵⁾ http://www.medizin.uni-koeln.de/kai/igmg/studien/ebm/ebm_pdf/ebm.pdf

⁶⁾ Hönel, A. **Vom soft law zur Rechtsvorschrift** <http://www.aventis.at/med/diplomarbeitjus.pdf>

⁷⁾ http://www.medizin-2000.de/streitpunkt/texte/bayer_kritik_hps.html

⁸⁾ Cochrane AL. **Random reflections on health services: Effectiveness and efficiency.** 1971 (reprinted by RSM Press 1999).

⁹⁾ **Anmerkung:** Ich bitte um Zusendung von Literatur, sofern meine Recherche in diesem Punkt fehlerhaft war!

¹⁰⁾ Mit dem Begriff **Goldstandard** taucht ein weiterer Ausdruck auf, der vermeintlich bekannt ist, aber in Wahrheit aus einem nicht für medizinische Studien relevanten Bereich stammt, nämlich dem Finanzwesen. Dort beschreibt „Goldstandard“ die verschiedenen Arten der Währungsdeckung.

¹¹⁾ Gigerenzer, Gerd (2002). **Das Einmaleins der Skepsis.** Über den richtigen Umgang mit Zahlen und Risiken. Berlin: Berlin Verlag.

¹²⁾ Krämer, Walter (1991). **So lügt man mit Statistik.** Frankfurt/M.: Campus.

¹³⁾ Der „**Placeboeffekt**“ hat einen negativen Beigeschmack. Allerdings gibt es wissenschaftlich angewandte Formen des Placebo im assertiven Training, bei der Konditionierung oder der Entwöhnung, die längst zu anerkannten Therapien geworden sind.

¹⁴⁾ http://www.medizin-2000.de/streitpunkt/texte/bayer_kritik_polypill.html