

Von der Eminenz- zur Evidence-basierten Medizin

Prof. em. Dr. med. Johann Steurer

Zunächst ist eine begriffliche Klärung notwendig: Im Titel wird bewusst der Begriff «Evidence» und nicht «Evidenz» verwendet. Während «Evidenz» laut Duden etwas als offensichtlich, klar, plausibel und schlüssig beschreibt, bezeichnet «Evidence» im englischen Sprachgebrauch «something helpful in forming a conclusion or judgement» und ist damit weniger eindeutig und normativ. Diese Unterscheidung ist wesentlich für das Verständnis der evidenzbasierten Medizin (EBM), insbesondere im Hinblick auf die Auseinandersetzung mit dem Vorwurf, sie reduziere ärztliches Handeln auf „Kochbuchmedizin“.

Faktoren und Rahmenbedingungen der Entstehung der EBM.

Im Wesentlichen trugen drei Faktoren zur Entstehung der EBM bei: Erstens nahm der uneingeschränkte Glaube an die Aussagen medizinischer Autoritäten zunehmend ab. Zweitens wurden immer mehr Daten zur erheblichen Variabilität in der Behandlung von Patientinnen und Patienten publiziert. Drittens fehlte vielen Medizinerinnen und Medizinern die Ausbildung, um die stetig wachsende Zahl an Studien zu Diagnostik, Therapie und anderen medizinischen Fragestellungen kritisch zu beurteilen.

Der Glaube an die Aussagen medizinischer Autoritäten wurde in Frage gestellt

Vor 50 und mehr Jahren war medizinisches Wissen in hohem Mass durch die persönlichen Erfahrungen und Einschätzungen angesehener Ärzte geprägt. Ihren Empfehlungen, die sich auf den damaligen Stand der Wissenschaft – vorwiegend Pathologie und Pathophysiologie – und eine langjährige klinische Erfahrung stützten, wurde mit Respekt begegnet und in der Praxis mehr oder weniger umgesetzt.

Metaanalysen oder evidenzbasierte Leitlinien, wie sie heute als Referenz für medizinisches Wissen dienen, existierten zu jener Zeit noch nicht. Zudem waren nur wenige Studien publiziert, die den heute etablierten methodischen Standards entsprachen. So erschien die erste Arbeit, welche belegen konnte, dass eine Blutdrucksenkung bei Hypertonikern das Risiko für Herzinfarkte reduziert, erst 1970 (1).

Der Nachweis, dass durch den Einsatz eines cholesterinsenkenden Medikaments, eines Statins, die Herzinfarkttrate gesenkt werden kann, gelang sogar erst deutlich später: Die vier S-Studie wurde 1994 publiziert (2).

Diese Erkenntnisse zur positiven Wirkung der Blutdruck- und Cholesterinsenkung mussten zunächst innerhalb der Ärzteschaft verbreitet werden. Dabei spielten medizinische Autoritäten – häufig als *Opinion Leader* bezeichnet – eine wichtige Rolle. Medien in den 1980er-Jahren (aber auch noch später) berichteten, dass einzelne dieser *Opinion Leaders* – nicht alle – an der Weitergabe solcher Botschaften sehr gut verdienten. Dies liess Zweifel an ihrer Unabhängigkeit bei der Beurteilung der Wirksamkeit aufkommen und schwächte das Vertrauen in sie. Wiederholt wurde über enge Verbindungen zwischen der pharmazeutischen Industrie und medizinischen Meinungsführern berichtet (3). Besonders aufmerksam verfolgt wurden entsprechende Hinweise in mehreren Publikationen von Marcia Angell, der späteren Chefredaktorin des *New England Journal of Medicine* (4, 5)¹. Auch Boulevardmedien griffen das Thema auf, etwa mit Schlagzeilen wie: «*Professor lobt Todespille und sahnt ab*».

Zu jener Zeit herrschte nur geringe Transparenz darüber, in welchem Ausmass Ärzte finanzielle Zuwendungen von der Pharmaindustrie oder anderen medizinnahen Branchen erhielten. Zwar wird heute häufig von Transparenz gesprochen, doch ob diese tatsächlich umfassend gewährleistet ist, bleibt offen. Unbestritten ist hingegen, dass solche Enthüllungen das Vertrauen in die Medizin beeinträchtigt haben: Viele Menschen waren und sind sich nicht mehr sicher, ob ihre Interessen im Zentrum standen – oder ob andere, finanziell attraktivere Motive überwogen.

Die medizinisch nicht erklärbare Variabilität in der Behandlung von Patienten wurde offensichtlich.

Bereits vor mehr als 50 Jahren publizierte John Wennberg eine Arbeit (6) – es folgten noch mehrere Dutzende Publikationen zu dem Thema – in denen er zeigte, dass z.B. die Anzahl Operationen an der Wirbelsäule an den Koronararterien und der Prostata zwischen einzelnen geographischen Bezirken in den USA sehr unterschiedlich waren.

¹ Teilweise betreffen die angeführten Referenzen Publikationen, die nach Einführung der EBM publiziert wurden; dennoch verweisen sie auf Probleme, die der Einführung der EBM zeitlich vorausgingen.

Kolektomien und die operative Versorgung von Schenkelhalsfrakturen unterschieden sich aber kaum zwischen den einzelnen Bezirken. Die Indikation für die Operation der Schenkelhalsfraktur und die Kolektomie wegen eines Karzinoms sind klar. Die Indikationen für eine Rückenoperation eine Bypassoperation oder eine Prostataoperation (Hyperplasie) sind dehnbar. Die teilweise hohen Zahlen an Operationen sind Hinweise auf eine Überversorgung, also Operationen, die medizinisch nicht nötig wären.

Auch aus der Schweiz gab es eine Studie, in der die Variabilität der medizinischen Behandlung vor und nach einer Medienkampagne gezeigt wurde. Domenighetti publizierte 1988 (7) über die Häufigkeiten von Hysterektomien in den Kantonen Tessin und Bern. Nach einer Kampagne im Tessin mit dem Ziel die Indikation zur Hysterektomie etwas strenger zu stellen, nahm die Rate an Hysterektomien um mehr als zwanzig Prozent ab, in Bern stieg die Rate dagegen leicht an.

Fehlende Kompetenzen für die kritische Beurteilung der Publikationen, deren Zahl zudem exponentiell anstieg

Nachdem die Zulassungsbehörden (FDA und andere) die Zulassungskriterien für Medikamente verschärften nahm die Anzahl publizierter Studien seit den 1960iger Jahren kontinuierlich zu. Die Anzahl randomisierter Studien, die einen kleinen Teil der gesamten Publikationen ausmachen, stieg gemäss PubMed² von 22 Publikationen (1950 bis 1959) auf 1281 (1960 bis 1969) und auf 56'000 (1980 bis 1989). Eine 2010 erschienene Arbeit berichtete, dass zum damaligen Zeitpunkt täglich 75 randomisierte Studien und 11 Systematische Übersichten (systematic Reviews, SR) publiziert wurden (8).

Für Ärztinnen und Ärzte wurde es zunehmend schwieriger – eigentlich unmöglich –, sich in der stetig wachsenden Flut medizinischer Publikationen zurechtzufinden. Zudem fehlte es damals an einer systematischen Ausbildung in der kritischen Bewertung wissenschaftlicher Studien.

An der neu gegründeten McMaster University in Ontario, Kanada, setzten sich mehrere Professoren intensiv mit diesem Problem auseinander. Sie gelangten zu der

² PubMed ist eine Datenbank mit medizinischen Publikationen bezogen auf den gesamten Bereich der Biomedizin.

Erkenntnis, dass Medizinerinnen und Mediziner gezielt in der kritischen Analyse wissenschaftlicher Literatur geschult werden müssten.

Der Begriff «Evidence-based Medicine» erschien erstmals 1991 in einer wissenschaftlichen Veröffentlichung (9). Bereits ein Jahr später publizierte eine Gruppe von Ärzten der McMaster University einen Artikel im *Journal of the American Medical Association*, der die medizinische Praxis weltweit nachhaltig prägen wird. Der Titel dieser Arbeit lautete: „Evidence-based Medicine: A new Approach to Teaching the Practice of Medicine“ (10).

Wie ist EBM definiert und wie soll dieses Konzept in der Praxis umgesetzt werden?

Im ersten Absatz der JAMA-Publikation betonen die Autoren, dass bei klinischen Entscheidungen der Intuition, der klinischen Erfahrung und pathophysiologischen Überlegungen ein geringeres Gewicht zukommen soll als den Ergebnissen klinischer Studien. In einer später von D. Sackett publizierten Definition der EBM wird der klinischen Erfahrung jedoch wieder ein höherer Stellenwert eingeräumt (11). Innerhalb der EBM-Gemeinschaft bleibt die „klinische Erfahrung“ dennoch weitgehend eine Black Box, zu der bislang nur wenig geforscht und publiziert wurde. Klinische Erfahrung wird immer wieder am Rande erwähnt, aber was das genau ist, blieb unklar.

Die Definition von Sackett lautet.

«EBM ist der sorgfältige, klare und wohl überlegte Gebrauch der besten Evidence bei Entscheidungen für die Behandlung von Patienten. Das Praktizieren der EBM bedeutet die Integration der individuellen Expertise mit in der Forschung generierter bester Evidenz» (11).

Zur Umsetzung des Konzeptes, den Ergebnissen der klinischen Forschung bei medizinischen Entscheidungen einen höheren Stellenwert beizumessen, schlugen die Autoren folgendes vor.

1. Wenn ein klinisches Problem zu lösen ist, wird in einem ersten Schritt eine "beantwortbare" Frage formuliert. *Das Beispiel in der ersten Publikation von 1992 war; ein 43-jähriger Mann hatte einen ersten epileptischen Anfall. Die Abklärungen mit Bildgebung zeigten keine Veränderungen, die den Anfall erklären würden. Die Frage des Patienten war, wie wahrscheinlich er mit einem zweiten epileptischen Anfall zu rechnen hat?*

2. In einem zweiten Schritt wird nach Studien gesucht, welche diese in Schritt 1 gestellte Frage (hoffentlich) beantworten kann.
3. In einem dritten Schritt wird die identifizierte Studie (oder Studien) gelesen und kritisch beurteilt (kurz; wie glaubwürdig sind die Ergebnisse der Studie (n)).
4. Die Ergebnisse der Studie werden mit dem Patienten besprochen und dann eine Entscheidung getroffen

In der Theorie erscheint dieses Vorgehen überzeugend, in der klinischen Praxis ist es jedoch aus zeitlichen und fachlichen Gründen kaum realisierbar. Die kritische Beurteilung klinischer Studien ist anspruchsvoll und setzt fundierte Kenntnisse der Studiendesigns voraus. Darüber hinaus erfordert die Bewertung der korrekten Anwendung statistischer Methoden spezialisiertes methodisches Wissen, das ein eigenes Themenfeld darstellt.

Nachdem sich diese erste Phase der Evidence-based Medicine als nur begrenzt praktikabel erwiesen hatte, wurden in einer weiteren Entwicklungsstufe Methoden eingeführt und kontinuierlich verfeinert, mit denen die Ergebnisse mehrerer Studien zu einer spezifischen Fragestellung systematisch zusammengeführt werden. Dieses Vorgehen wird als «systematischer Review» mit oder ohne Metaanalyse bezeichnet. Ein kürzlich publizierter systematischer Review ohne Metaanalyse befasste sich beispielsweise mit der Frage, ob bei Patientinnen und Patienten mit Herzinsuffizienz eine Einschränkung der Trinkmenge empfohlen werden sollte (12).

Guidelines sind eine weiterführende Verdichtungsstufe wissenschaftlicher Forschungsergebnisse. Grundlage ihrer Erstellung sind Originalstudien, systematische Reviews mit oder ohne Metaanalysen sowie die klinische Expertise von Fachpersonen, aus denen konkrete Empfehlungen abgeleitet werden. Ein Beispiel sind die Guidelines der Europäischen Gesellschaft für Kardiologie zur Behandlung von Patienten mit einem akuten koronaren Syndrom – ein Dokument mit 107 Seiten (13).

Was sind die intendierten und was die nicht intendierten Effekte der EBM?

Intendierte Effekte

Die immer wieder gestellte Frage, ob die Patientenbehandlung mit der EBM besser geworden sei, ist nicht beantwortbar. Es wird nie eine Studie geben, in der die Patientenbehandlung mit EBM versus non-EBM verglichen wird.

Unbestritten ist, dass die Medizinerinnen und Mediziner nach Abschluss des Medizinstudiums ein besseres Verständnis über die Vor- und Nachteile verschiedener Studiendesigns haben, als dies vor 40 Jahren der Fall war. Es ist allgemein bekannt, dass bei der Untersuchung der Wirksamkeit eines Medikamentes oder einer anderen Intervention eine sorgfältig durchgeführte randomisierte Studie vertrauenswürdiger Ergebnisse liefert als eine Fall-Kontroll-Studie.

Heute kennen die Studierenden den Unterschied zwischen einer absoluten und einer relativen Risikoreduktion und können erklären, was ein 95%-Konfidenzintervall ist.

Ein weiterer positiver Effekt ist die Entwicklung von Standards für die klinische Forschung (z.B. Struktur eines Forschungsprotokolls) und Standards für die Publikation von Studienergebnissen (z.B. CONSORT (14)) und Standards für die Formulierung von Empfehlungen (GRADE (15)).

Mehrere «systematic Reviews» beeinflussten und beeinflussen die klinische Praxis unmittelbar, beispielsweise durch den Nachweis, dass die Bauchlage im Schlaf das Risiko des plötzlichen Kindstods senkt, oder dass eine zusätzliche Steroidtherapie neben Antibiotika den Outcome von Patienten mit Lungenentzündung verbessert (16, 17).

Ob das vermehrte und kritischere Hinterfragen therapeutischer Maßnahmen eine direkte Folge der EBM ist, lässt sich nicht eindeutig beantworten. Eine Behandlung ohne belastbare Evidence, wie sie beispielsweise beim Einsatz von Lidocain nach Myokardinfarkt erfolgte, wäre heute kaum mehr vorstellbar. Rückblickend zeigte sich, dass die prophylaktische Gabe von Lidocain zur Vermeidung schwerer Rhythmusstörungen nach einem Herzinfarkt mit einer erhöhten Mortalität einherging (18).

Nicht-intendierte Konsequenzen und Kritik an der EBM

Die Bezeichnung der EBM als «Kochbuchmedizin» – mit dem Vorwurf ist die EBM konfrontiert – geht von der Annahme aus, dass Studienergebnisse medizinische Entscheidungen eindeutig determinieren. Diese Sichtweise resultiert aus der deutschen Bedeutung des Begriffs Evidenz. Dass Evidenz jedoch nicht mit Entscheidungszwang gleichzusetzen ist, sollte allen Medizinern klar sein: Therapieentscheidungen entstehen aus der Abwägung wissenschaftlicher Ergebnisse mit klinischer Erfahrung sowie den Präferenzen und Wünschen der Patienten.

Ein weiterer Kritikpunkt ist die Tendenz zur **Trivialisierung** klinischer Forschung. Die hierarchische Darstellung der Evidenz in Form einer Pyramide mit «systematic Reviews» an der Spitze und randomisierten Studien auf der nachgeordneten Ebene suggeriert eine Bewertungslogik, die in dieser Form nicht haltbar ist. Insbesondere lässt sich nicht rechtfertigen, weshalb ein «systematic Review» aus methodisch limitierten Studien als glaubwürdiger gelten soll als eine einzelne, sorgfältig durchgeführte Studie. Ebenso stellen generalisierende Aussagen, wonach Effekte von Screening-Massnahmen ausschließlich in randomisierten Studien untersucht werden könnten, eine problematische Vereinfachung dar. Das häufig vorgebrachte Argument, dass ohne randomisierte Studien keine belastbaren Aussagen möglich seien, greift zu kurz. In Disziplinen wie der Ernährungsmedizin mangelt es häufig an solchen Studien. Dies schließt eine «evidence-basierte» Entscheidungsfindung jedoch keineswegs aus, da sich auch aus Beobachtungsdaten und pathophysiologischen Modellen relevante Erkenntnisse gewinnen lassen.

Zukunft

Welche der drei eingangs genannten Faktoren – der nicht unberechtigt schwindende Glauben an Autoritäten, die hohe Variabilität in der Patientenbehandlung sowie die fehlenden Kompetenzen zur kritischen Beurteilung der Literatur – die zur Entstehung der evidenzbasierten Medizin beigetragen haben, sind heute noch von Bedeutung? Aussagen von Chefärzten wie „Das haben wir schon immer so gemacht“ hört man heute deutlich seltener als noch vor 50 Jahren. Verbindungen zur Pharmaindustrie werden inzwischen transparent offengelegt; welche Effekte diese Verflechtungen tatsächlich immer noch haben, ist aber unklar.

Ob die Variabilität der Behandlung von Patienten abgenommen hat, lässt sich nicht generell beantworten. Schaut man sich die Verschreibungsrate von ADHS-Medikamenten in der Schweiz an, dann sind doch beträchtliche Unterschiede zwischen den Kantonen feststellbar (19). Allem Anschein nach wird die Verschreibungsrate von Medikamenten – nicht nur bei ADHS – sowie die Operationsindikation nicht ausschließlich durch die Empfehlungen der Guidelines bestimmt.

Heute sind Studierende nach dem Studium besser in der Lage Publikationen kritisch zu beurteilen, als dies vor 40 Jahren der Fall war.

Die Frage, die man sich heute stellen muss, ist die nach dem Umgang mit der enormen Zunahme an wissenschaftlichen Publikationen. Es ist heute unmöglich sich in den Literaturbergen zurechtzufinden.

Im Jahr 2025 wurden gemäss PubMed die Ergebnisse von rund 50 000 randomisierten Studien sowie 36 609 systematischen Reviews publiziert – das entspricht etwa 100 systematischen Reviews pro Tag. Nur ein minimster Teil dieser Arbeiten (manche sagen weit weniger als 1%) ist von ausreichender Relevanz, um Eingang in klinische Guidelines zu finden (20-22). Zudem kam Ioannidis zum Schluss, dass rund 80 % der medizinischen Publikationen entweder falsch, verzerrt und/oder klinisch irrelevant sind (23).

Vor diesem Hintergrund wird deutlich, dass «jemand» die Ergebnisse der publizierten Studien systematisch synthetisieren und das daraus gewonnene Wissen so aufbereiten sollte, dass es in der klinischen Praxis anwendbar ist – idealerweise nicht nur auf Populationsebene, sondern für den individuellen Patienten.

Selbst Guidelines als derzeit höchstverdichtete Form medizinischer Information lösen das Informationsproblem nicht. In einer Notfallstation wurden konsekutiv 18 Patienten mit insgesamt 44 Diagnosen erfasst und die dazu verfügbaren Guidelines gesucht. Die zu diesen Erkrankungen verfügbaren Guidelines umfassten zusammen 3 670 Seiten (24). Es ist illusorisch anzunehmen, dass diese Informationsmenge im klinischen Alltag in angemessenem Mass rezipiert werden kann. Möglicherweise kann die künstliche Intelligenz künftig helfen, relevante Informationen in diesen nahezu 4 000 Seiten gezielt zu identifizieren.

Ein Forscher, der sich über Jahre intensiv mit dem Problem der Evidenzsynthese beschäftigt hat, formulierte einen Vorschlag zur Lösung des Problems. Derzeit hängt

die akademische Qualifikation klinischer Professorinnen und Professoren primär von der Anzahl publizierter Originalarbeiten ab, wobei die Grundlagenforschung deutlich höher gewichtet wird als die angewandte klinische Forschung.

Wünschenswert wäre eine Neubewertung akademischer Exzellenz: Klinische Professorinnen und Professoren sollten Expertinnen und Experten der Synthese und Interpretation medizinischer Evidenz sein. Beherrschten sie diese Kunst auf höchstem Niveau, könnten sie mit Recht als „Eminenzen“ bezeichnet werden – und der Begriff der Eminenz würde in der Medizin seine negative Konnotation verlieren (25).

References

1. Effects of treatment on morbidity in hypertension. II. Results in patients with diastolic blood pressure averaging 90 through 114 mm Hg. *Jama*. 1970;213(7):1143–52.
2. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet*. 1994;344(8934):1383–9.
3. Weiss H. *Korrupte Medizin. Ärzte als Komplizen der Konzerne.*: Kiepenheuer & Witsch; 2008.
4. Angell M. *Drug Companies & Doctors: A Story of Corruption*. New York Review of Books. 2009 January 15.
5. Angell M. Is Academic Medicine for Sale? *NEJM*. 2000;342(20):1516–8.
6. Wennberg J, Gittelsohn. Small area variations in health care delivery. *Science*. 1973;182(4117):1102–8.
7. Domenighetti G, Luraschi P, Gutzwiller F, Pedrinis E, Casabianca A, Spinelli A, et al. Effect of information campaign by the mass media on hysterectomy rates. *The Lancet*. 1988;332(8626-8627):1470–3.
8. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS medicine*. 2010;7(9):e1000326.
9. Guyatt GH. Evidence-Based Medicine. *Annals of Internal Medicine*. 1991;114:A–16.
10. Guyatt G, Cairns J, Churchill D, Cook D, Haynes B, Hirsh J, et al. Evidence-based medicine: a new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420–5.

11. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71–2.
12. Herrmann JJ, Brunner-La Rocca HP, Baltussen L, Beckers-Wesche F, Bekkers S, Bellersen L, et al. Liberal fluid intake versus fluid restriction in chronic heart failure: a randomized clinical trial. *Nat Med*. 2025.
13. ESC Clinical Practice Guidelines. 2023 ESC Guidelines for the management of acute coronary syndromes 2023 [Available from: <https://www.escardio.org/Guidelines/Clinical-Practice-Guidelines/Acute-Coronary-Syndromes-ACS-Guidelines>].
14. Hopewell S, Chan A-W, Collins GS, Hróbjartsson A, Moher D, Schulz KF, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. *The Lancet*. 2025;405(10489):1633–40.
15. Guyatt G, Agoritsas T, Brignardello-Petersen R, Mustafa RA, Rylance J, Foroutan F, et al. Core GRADE 1: overview of the Core GRADE approach. *Bmj*. 2025;389.
16. Gilbert R, Salanti G, Harden M, See S. Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. *Int J Epidemiol*. 2005;34(4):874–87.
17. Siemieniuk RA, Meade MO, Alonso-Coello P, Briel M, Evaniew N, Prasad M, et al. Corticosteroid Therapy for Patients Hospitalized With Community-Acquired Pneumonia: A Systematic Review and Meta-analysis. *Ann Intern Med*. 2015;163(7):519–28.
18. Moore JM. *Deadly Medicine. Why tens of thousands of heart patients died in America's worst drug disaster*: Simon & Schuster; 1995.
19. Schweizerisches Gesundheitsobservatorium. ADHS Medikament 2025 [Available from: <https://www.versorgungsatlas.ch/indicator/ 042>].
20. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet*. 2014;383(9912):166–75.
21. Ioannidis JP. Why Most Clinical Research Is Not Useful. *PLoS Med*. 2016;13(6):e1002049.
22. Moher D, Glasziou P, Chalmers I, Nasser M, Bossuyt PMM, Korevaar DA, et al. Increasing value and reducing waste in biomedical research: who's listening? *Lancet*. 2016;387(10027):1573–86.
23. Ioannidis JP. Why most published research findings are false. *PLoS medicine*. 2005;2(8):e124.

24. Allen D, Harkins KJ. Too much guidance? *Lancet*. 2005;365(9473):1768.
25. Miettinen OS. Evidence in medicine: invited commentary. *CMAJ*. 1998;158(2):215–21.