

Dr. med. Johannes G. Schmidt, Allgemeinpraxis & Institut für Klinische Epidemiologie, Einsiedeln

Evidence-Based Medicine und der «Publication Bias»

Der sogenannte Publication Bias entsteht dadurch, dass mit einer gewissen Regelmässigkeit Studien mit negativen Ergebnissen in den Schubladen verschwinden, solche mit positiven Ergebnissen aber zur Publikation gelangen. Neuerdings wird auch noch von einem Einfluss der Publikationssprache auf die Berücksichtigung in Meta-Analysen berichtet und der Begriff eines Turm-zu-Babel-Bias kreiert. Wie wichtig sind aber Meta-Analysen und ihre Fehlermöglichkeiten für eine auf Evidenz gestützte Praxis, und stehen andere Publikations-Misereen für den praktischen Anwender vielleicht eher im Vordergrund?

Evidenz und klinisches Ermessen in der Praxis

Evidence-Based Medicine ist im Grunde ein neues Schlagwort für die Disziplin der klinischen Epidemiologie. Diese hat schon immer eine kritische Interpretation vorliegender Studienergebnisse nach praxisgerechten Beurteilungsgrössen gelehrt sowie eine sorgfältige Planung und Durchführung klinischer Studien angestrebt mit dem Ziel praktisch aussagekräftiger und relevanter Studienergebnisse. Die kunstgerechte, kritische Interpretation («critical appraisal») ist indessen ein Werkzeug, welches das klinische Ermessen («judgement») nicht ersetzen kann. Allerdings ist ohne kunstgerechtes Zerpflücken der vorliegenden Evidenz auch kein vernünftiges Judgement möglich. Die Ergebnisse guter und richtig interpretierter Studien sind somit für die Praxis notwendig, aber nicht hinreichend. Es braucht ein Abwägen der vorhandenen Evidenz mit den Vorlieben und Möglichkeiten des Patienten und den eigenen praktischen Erfahrungen des Arztes. Dieser Sachverhalt wird auch von den nach wie vor klinisch tätigen Autoren des Begriffs Evidence-Based Medicine ausdrücklich festgehalten, und der klinische Epidemiologe Feinstein etwa ist dem Begriff Evidence-Based Medicine gegenüber skeptisch, weil er eine Verarmung und Reduktion der Medizin auf die einfachen und einseitigen Fragen befürchtet, für welche Evidenz vorliegt. Es wäre deshalb völlig falsch zu meinen, mit Evidence-Based Medicine seien unsere Probleme in der Praxis schon gelöst, endlich seien einfache Antworten für komplizierte Patientenprobleme da. Mindestens in der hausärztlichen Praxis führt Evidence-Based

Medicine oft eher dazu, dass man bei vielen Patientenproblemen nun überhaupt nichts mehr Rechtes weiss, weil sich das bisher in Weiter- und Fortbildungen Gelernte als Lehrmeinung oder Glaube entpuppt, der gar nicht durch aussagekräftige Evidenz substantiiert werden kann. Dies hat neben natürlichen methodischen Schwierigkeiten damit zu tun, dass die wichtige Aufgabe der klinischen Forschung, Studien durch einen Kompromiss zwischen Spezifität der Therapie und Praxisnähe ihrer Anwendung erst aussagekräftig zu machen (vgl. «Evidence-Based Medicine und «Placebo», SÄZ Nr. 42/97), aus Unbedachtheit und aufgrund partikularer Interessen vernachlässigt wird. Zudem ist die Evidenz oft auf die Behandlung paraklinischer Parameter beschränkt; wir müssen dann zweifeln, ob etwa die erfolgreiche EKG-Behandlung, Laborwertverbesserung oder anatomische Rekonstruktion dem Patienten wirklich nützt. So haben wir heute – um ein Bonmot unter Allgemeinpraktikern aufzunehmen – viele Antworten (Studienergebnisse) auf Fragen (Patientenprobleme), die wir so gar nicht haben. Evidence-Based Medicine wird deshalb oft auch mit einigem Recht als «Befreiungstheologie» für die Praxis bezeichnet, weil die sich offenbarende institutionelle Ignoranz dem eigenen Erkennen und Verstehen neuen Raum lässt. Freiheit indessen verpflichtet, und gerade deshalb sind die Denkschulung und die Methodik der Evidence-Based Medicine eine Voraussetzung, um die vorhandenen Probleme und Lücken explizit zu machen und zu formulieren. Um Missverständnissen vorzubeugen, möchte ich deshalb festhalten, dass aus der vorhandenen Misere nicht ein Alleinkämpfertum in der Praxis abgeleitet werden kann, welches sich einer Erklärung bzw. kritischen Beurteilung seiner «Kunst» verweigert. Andererseits gilt es aber mit der alten Praktiker-Tugend im modernen Kleid der klinischen Epidemiologie zu durchschauen und zu quantifizieren, welches Wissen wirklich ein Gewicht hat oder vernachlässigt werden kann.

Meta-Analysen sind ein Randproblem

Es soll nicht ohne Zufriedenheit vermerkt werden, dass die FMH ebenfalls diesen Weg erkannt hat, wenn sie Kollege von Below im Newsletter der Cochrane-Arbeitsgruppe Schweiz schreiben lässt: «Statt von allen

Seiten mit statistisch hochsignifikanten neuen Erkenntnissen bombardiert zu werden, ermuntert die – lernbare – Methode des Critical Appraisal und das Verständnis der Prinzipien der besten externen Evidenz dazu, sich eine eigene Meinung zu bilden. Evidence-Based Medicine will nicht Instrument zu einer «Kochbuchmedizin» sein ...»

Bei dieser Aufgabe sind nun methodische Probleme von Meta-Analysen ein Nebenproblem, was nicht ausschliesst, dass diese bei speziellen Fragen gelegentlich ein ausschlaggebendes Gewicht haben. Meta-Analysen können im Grunde lediglich die statistische Signifikanz (den Power) erhöhen, können aber mangelnde klinische Relevanz nie korrigieren und sind auch in bezug auf ihre angeblich zuverlässigere Treffsicherheit und Generalisierbarkeit ein eher unsicheres Instrument. Dies wurde kürzlich im New England Journal of Medicine gezeigt: Für einen Drittel der sorgfältig analysierten Behandlungen (z. B. Nitroglyzerin bei Infarkt, Aspirin bei intrauteriner Wachstumsretardierung) zeigten vorgängige Meta-Analysen Resultate, die in nachfolgenden grossen Studien widerlegt worden waren. Es kann nicht geklärt werden, ob für solche Diskrepanzen veränderte Rahmenbedingungen über die Zeit, Unvergleichbarkeit der Patientengruppen aufgrund unbekannter Faktoren oder ein Publication Bias verantwortlich ist. Damit soll noch einmal gesagt sein, dass Evidence-Based Medicine in erster Linie eine Beurteilungsmethode des Zu-Ende-Denkens darstellt, welche Denkfehler und Trugschlüsse in herkömmlichen Lehrmeinungen erkennen, Wissen und Unwissen näher eingrenzen und die Wahrscheinlichkeit von Vor- und Nachteilen einer Intervention schätzen lässt. Dabei ist eine patientenrelevante Wirkgrösse in Form einer genügend niedrigen Number needed to treat eine zentrale Grösse (Number needed to treat: Wie viele Patienten muss ich den Auswirkungen einer Intervention aussetzen, um einmal einen Erfolg zu erzielen, bzw. mit welcher Wahrscheinlichkeit kann mein Patient einen Nutzen erwarten, mit 1 zu 10, 1 zu 100, 1 zu 1000?). Meta-Analysen und ihre Probleme mit dem Publication Bias sind neben diesen zentraleren Fragen praktisch von untergeordneter Bedeutung. Jedenfalls kann Evidence-Based Medicine keineswegs mit Meta-Analysen gleichgesetzt werden, wie manchmal der Anschein erweckt wird. So wie es aussieht, hängen solche dezentrale Prioritätensetzungen mit institutionellen Voraussetzungen ab, indem (im deutschsprachigen Raum) oft Institute mit Schreibtisch-Medizinern die Diskussion prägen. Was wir hierzulande deshalb vorrangig fördern müssen, ist die entsprechende Ausbildung praktisch-klinisch tätiger Kolleginnen und Kollegen und die schnelle Schaffung und Alimentierung von (dezentralen) Institutionen für klinische Epidemiologie, welche vorwiegend von praktisch und klinisch tätigen Fachleuten besetzt sind.

Die entscheidende und grosse Aufgabe der Cochrane Collaboration, der man Erfolg wünschen muss, ist die vollständige Registrierung von Studien, deren Durchführung läuft oder geplant ist. Weit effizienter als die Suche nach verlorenen, nicht publizierten Studien ist eine Konzentration der heutigen und künftigen Forschungsanstrengungen auf methodisch gute Studien, die wirklich aussagekräftig konzipiert und praktisch relevant sind. Dieses oft gedankenlose Daherpublizieren, weil man ohne angeblich untergehe («publish or perish»), sollte von der Ärzteschaft gestoppt werden; dann muss auch nicht mehr so viel Literatur, die von vornherein nichts taugt, kritisch evaluiert werden. Die Ressourcen könnten dann für weniger, dafür intellektuell angemessene Studien verwendet werden, bei denen die Mühe gemacht worden ist, eine anständige Forschungsfrage zu formulieren. Die Forschungsförderung hat in Zukunft vielmehr die Erarbeitung von Forschungsfragen zu unterstützen statt nur das Erheben von neuen Daten (die ohne adäquate Forschungsfrage oft wertlos sind). Der Publikation irrelevanter Studien voraus geht besonders hierzulande ein «Institutions-Bias», der sich unangenehm von dem weit offeneren und qualitätsorientierten akademischen Klima der angelsächsischen Länder unterscheidet. (Hier ist auch der neuralgische Punkt, welcher das derzeitige Unbehagen der immer noch betriebsblinden Schweiz zu einem guten Teil bestimmt.)

Effektive «Publikations-Biases»

Wichtiger als das Nicht-Publizieren ungünstiger Studienergebnisse dürfte schliesslich sein, wie Publikationen zustande kommen. Auch wenn ich keine systematische Übersicht über die Voreingenommenheit und Einflussnahme von Redaktionen kenne, so glaube ich, dass folgende zwei Beispiele aus der eigenen Karriere nicht Einzelfälle sind.

Bei unserer Analyse der Auswirkungen des Ultraschall-Screenings in der Schwangerschaft kamen Heiner Bucher und ich zum Schluss, dass weder Morbidität noch Mortalität verbessert werden, dass der Ultraschall aber als Missbildungs-Screening funktioniere. Diese Funktion des Routine-Ultraschalls sollte deshalb explizit gemacht werden, des weiteren sei darüber aufzuklären, dass bei diesem Vorhaben auch falsche Missbildungsdiagnosen und der Abort gesunder Kinder in Kauf zu nehmen seien. Erst weitere Studien könnten klären, ob der Nutzen der Missbildungs-Erkennung im Stadium der Abort-Möglichkeit nicht durch das Ausmass falsch positiver Diagnosen in Frage gestellt würde (unter der Voraussetzung, dass man sich dieser ethisch problematischen Zielsetzung überhaupt anschliesst). Diese Arbeit wurde im British Medical Journal erst publiziert, nachdem wir unsere Feststellungen auf Verlangen der Redaktion änderten und nun zu schreiben hatten, dass

der Routine-Ultraschall für viele Frauen von beträchtlichem Nutzen sei als Mittel zur Entdeckung von Missbildungen. Diese Diktion wurde in der Folge dann auch häufig zitiert. Aus der möglichen Funktion als Missbildungs-Screening mit der inhärenten Problematik falsch positiver Diagnosen entstand auf Druck der Redaktion ein Nutzen als Missbildungs-Screening. Damit wurde eine weltanschauliche oder Wertungs-Frage, die nicht medizinisch zu beantworten ist, zu einer medizinischen Aussage umfabriziert, und die dringend nötige Transparenz zwischen Evidenz und Wertung verwischt – ein wirksamer «Publication Bias», der vermutlich oft eine Rolle spielt.

Beim zweiten Beispiel geht es um eine geradezu grotesk schiefgewinkelte Publikation wider besseres Wissen, welche in einer Schweizer Fachzeitschrift erfolgte. Das Beispiel ist besonders schön, weil die gleiche Arbeit später in einem internationalen Journal mit ganz anderen Schlussfolgerungen noch einmal veröffentlicht wurde. Das Beispiel ist zudem bezeichnend, als es aus dem Bereich der Sozial- und Präventivmedizin stammt, wo epidemiologisch-methodisches Wissen eigentlich zu erwarten wäre. Es handelt sich dabei um eine nordische «Studie» über die Wirkung der Brust-Selbstuntersuchung, welche an einer Schweizer Tagung vorgestellt wurde. Dabei wurde gezeigt, dass die Brust-Selbstuntersuchung die Brustkrebssterblichkeit um 30% gesenkt hätte, verglichen wurde eine mitmachbereite, motivierte Frauengruppe mit der Durchschnittsbevölkerung. Bereits im nächsten Vortrags-Dia war ersichtlich, dass auch die Gesamtsterblichkeit sich in den verglichenen Gruppen gleichsinnig um 30% unterschied; auch für methodische Anfänger ist damit ein sogenannter Selektions-Bias sofort ersichtlich (nicht die Therapie, sondern die Selektion der Studiengruppe ist für den Unterschied zur Kontrollgruppe verantwortlich). Die spätere Publikation in der Zeitschrift *Cancer* wies denn auch auf dieses gravierende Problem dieser Studie hin. Die entsprechende Schweizer Fachgesellschaft hingegen nahm die Studie zustimmend zur Kenntnis, ohne auf die elementaren methodischen Fehler überhaupt einzugehen. Der Publikation dieser Studie in der Schweizer Fachzeitschrift voraus gingen dann eine Korrespondenz und Telefonate, in welchen ich den entsprechenden Instituts-Vorsteher auf die Fehler dieser Studie und auf die Unhaltbarkeit der Schlussfolgerungen aufmerksam machte. Dennoch erfolgte die Veröffentlichung unter Aufrechterhaltung der Schlussfolgerung, die Brust-Selbstuntersuchung würde die Brustkrebsmortalität nachweislich um 30% senken, und unter Verheimlichung der «Ergebnisse», dass auch die Gesamtmortalität um 30% «gesenkt» wurde. Dabei war bereits aus anderen grossen, fehlerfreien Studien bekannt, dass die Brust-Selbstuntersuchung keinen Einfluss auf die Sterblichkeit hat, iatrogene Probleme aber zunehmen. Dieses

Beispiel zeigt, dass immer wieder institutionelle Interessen sogar wider besseres Wissen über die wissenschaftliche Redlichkeit gestellt werden, und es illustriert die beträchtliche Mühe, welche gerade auch die epidemiologisch geschulte Sozial- und Präventivmedizin mit einer klinischen Epidemiologie bzw. der Evidence-Based Medicine hat, wenn diese anfängt, Präventions-Dogmen zu hinterfragen. Wir wissen aus Gesprächen mit Exponenten der Evidence-Based Medicine aus Kanada, England und den USA, dass es auch dort der Präventions-Industrie im Verein mit der modernen Unfähigkeit, dem unvermeidlichen Sterben ins Auge zu sehen, immer noch gelingt, wissenschaftliche Evidenz herbeizureden, wo diese nicht vorhanden ist. Eine gewisse Ausnahme bildet der Schwangerschafts-Ultraschall, welcher in den USA und Kanada nicht empfohlen wird.

In der gegenwärtigen Glaubwürdigkeitskrise ist es für die Ärzteschaft als Ganzes aber nicht zu umgehen, diese Arten von institutionell bedingten «Publikations-Biases» allmählich zu korrigieren. □