DER ARZNEIMITTELBRIEF

Jg. 36, S. 89; Ausgabe 12 / 2002

Der Umgang mit der Evidenz

Der 18. Oktober 2001 – ein großer Tag in der Geschichte der Evidenz-basierten Medizin in Deutschland: Die deutsche Bundesregierung und die fünf im Bundestag vertretenen Parteien einigen sich, ein Deutsches Mammographie-Programm zur Früherkennung von Brustkrebs aufzulegen. Dabei berufen sie sich ausdrücklich auf die Evidenzlage aus sieben randomisierten Studien an 500000 Frauen. 10 Jahre nach Beginn der Diskussion über Evidenz-basierte Medizin ist die Leitidee ganz oben angekommen – ein großer Sieg für die Verfechter einer rationalen Medizin.

Der 18. Oktober 2001 – ein großer Tag in der Geschichte der Evidenz-basierten Medizin weltweit. Zeitgleich erschien die neue Version des Cochrane-Reviews über ”Screening for Breast Cancer with Mammography” (www.cochrane.de) und eine ausführliche Fassung des Reviews im Lancet (1). Mehr als 10 Jahre nach Geburt des Begriffs „EbM“ ist das regelmäßige Erstellen und „Updaten“ von Reviews Routine geworden. Das gleichzeitige Erscheinen beider Artikel bei der Cochrane Collaboration und in einem renommierten Journal ist außerdem ein Höhepunkt der wissenschaftlichen Diskurs-Kultur, hatten sich doch die Cochrane-Editoren mit einem Teil der Schlußfolgerungen der Reviewer nicht anfreunden können, den man deshalb jetzt nur im Lancet nachlesen kann. In den gemeinsamen Teilen und der diesbezüglichen Schlußfolgerung sind aber beide Versionen identisch. Welch eine kulturelle Errungenschaft: we agree where we agree and where we disagree.

Es schien also gut zu stehen um die Idee der Evidenz-basierten Medizin im Oktober 2001, wenn, ja wenn da nicht ein kleiner Schönheitsfehler wäre: Während die deutsche Politik unisono konstatiert, daß sich laut Evidenzlage die Brustkrebs-Mortalität von 50-70jährigen Frauen durch Mammographie-Programme um 20-30% senken lasse, stellen Cochrane und Lancet übereinstimmend fest: In diesem Review gelang es nicht, eine Reduktion der Gesamt-Mortalität festzustellen. Den qualitativ besten Studien gelang es darüber hinaus nicht einmal, eine Reduktion der Brustkrebs-Mortalität zu belegen.

Hohe Evidenz für zwei sich gegensätzlich ausschließende Positionen? Wie ist das möglich? Haben die Autoren verschiedene Studien herangezogen? Muß man das Kleingedruckte lesen, um die Aussagen richtig zu interpretieren? Gibt es gar eine dritte Möglichkeit zwischen Ja und Nein? Nichts dergleichen ist der Fall. Beide Urteile beruhen auf denselben sieben Studien, an beiden Bewertungen sind namhafte Wissenschaftler beteiligt, die sich auf die gleichen Prinzipien und Qualitätsmaßstäbe berufen. Eine Kompromiß-Position zwischen Ja und Nein gibt es nicht: Wenn die Mammographie nutzt, muß ein Vorsorgeprogramm mit großem persönlichem Engagement und finanziellem Aufwand umgesetzt werden, um die in Aussicht stehende Ernte auch wirklich in die Scheuer zu fahren. Hier sollte Evidenz-basierte Medizin die Grundlage liefern – und nun diese Divergenz in den Stellungnahmen. Kein Zweifel: die Evidenz-basierte Medizin hat Erklärungsnotstand.

Diese Publikation ist urheberrechtlich geschützt. Vervielfältigung sowie Einspeicherung und Verarbeitung in elektronischen Systemen ist nur mit Genehmigung der Herausgeber des ARZNEIMITTELBRIEFS gestattet.

Es ist wohl kein Zufall, daß dieses Desaster die Evidenz-basierte Medizin nicht bei der Bewertung von therapeutischen Zugängen, sondern in der Diagnostik ereilt, und hier insbesondere im Bereich der Vorsorge, der sogenannten Primär-Prävention bei symptomlosen augenscheinlich Gesunden. In der therapeutischen Forschung werden die Cochrane-Reviews trotz partikulärer Kontroversen im Allgemeinen sehr geschätzt. Sie geben Hilfestellungen in Entscheidungssituationen und Sicherheit im ärztlichen Alltag. In der Diagnostik hingegen ist das Beratungs-Bedürfnis auf Arzt- wie auf Patientenseite im Allgemeinen weniger stark ausgeprägt. Es erscheint doch so einfach: wir gehen auf Nummer Sicher und probieren einfach mal alles aus, was uns die moderne Medizinwelt so zu bieten hat. Information kann nie schaden, und wenn man für alles Vorsorge getroffen hat, kann man ruhiger schlafen, es wird schon nichts passieren, oder zumindest hat man alles getan, was in der eigenen Macht steht. Flächendeckende Vorsorge-Programme sind doch geradezu Pflicht im modernen Wohlfahrtsstaat.

Diese Philosophie der Maximal-Diagnostik würde auch tatsächlich wundervoll funktionieren (und ihre Grenze erst in der Kosten-Explosion des Gesundheitswesens finden), wenn die verwendeten Diagnose-Methoden tatsächlich so perfekt wären, wie wir sie uns wünschen, und – eine Zusatz-Bedingung, die oft übersehen wird – wenn eine wirksame Therapie für die jeweils zu diagnostizierende Krankheit vorliegt, die Nutzen aus dem durch das Screening gewonnenen Zeitvorsprung ziehen kann. Dabei wird im Allgemeinen völlig unterschätzt, welch hohe Anforderungen im jeweiligen Fall sowohl an die Güte der Diagnostik wie an die Effektivität der Therapie gestellt werden müssen. Die entscheidende Einflußgröße ist die Häufigkeit oder, besser gesagt, die Seltenheit der Erkrankung, gegen die Vorsorge getroffen werden soll. Wird z.B. ein Screening-Programm mit 99% Sensitivität und 99% Spezifität (das wäre eine sehr gute Quote) auf eine Population mit einer Erkrankungs-Prävalenz von 1 auf 1000 angewendet, so ist im Mittel 99% von 0,1% der Screening-Population krank und Test-positiv (Sensitivität), hingegen 1% der Population (100 minus Spezifität) falsch positiv (Spezifität = richtig Negative pro alle Gesunde). In unserem Falle kämen also auf einen richtig positiven 10 falsch positive Patienten. Wird auf diesen Personenkreis eine aufwendigere Diagnostik bzw. eine wirksame Therapie mit wenigen Nebenwirkungen angewandt, so mögen die Chancen noch recht gut sein, einen positiven Effekt zu erzielen. Wird hingegen dasselbe Screening-Programm auf eine Population mit einer Erkrankungs-Prävalenz von 1 auf 100000 angewendet, so befinden sich im Mittel unter 1000 Test-Positiven nur noch ein Kranker und 999 Falsch-Positive. Unter diesen Umständen hat es auch eine gute Therapie sehr schwer, überzeugende Ergebnisse zu erzielen.

Die Erfolgsbilanz eines solchen Programms wird weiter geschmälert, wenn die Qualität des Screening-Verfahrens in der Breitenanwendung nachläßt, das Screening zur Vorsicht auf unverdächtige Personengruppen ausgedehnt wird, die Screening-Intervalle unangemessen eng gewählt werden, andere Präventions-Maßnahmen vernachlässigt werden, die Therapie-Indikation zu spät (oder genauso verheerend: zu großzügig) gestellt wird oder die Nebenwirkungen nicht beachtet werden. In der Realität ist ein Vorsorge-Programm für seltene Erkrankungen deshalb alles andere als ein einfaches Nachgucken und auf Nummer-Sicher-Gehen. Der statistische Hebel ist grausam: jede auch nur geringfügige Reduktion der Prävalenz muß durch geradezu übermenschliche Anstrengungen kompensiert werden, um zum Erfolg zu führen. Es ist tatsächlich die sprichwörtliche Suche nach der

Diese Publikation ist urheberrechtlich geschützt. Vervielfältigung sowie Einspeicherung und Verarbeitung in elektronischen Systemen ist nur mit Genehmigung der Herausgeber des ARZNEIMITTELBRIEFS gestattet.

Nadel im Heuhaufen, die auch bei der theoretisch perfekten Nachforschung mit einem Magneten in der Praxis an der Vielzahl anderer metallischer Verunreinigungen unserer Umwelt scheitert. Und dabei ist in der Bilanz noch nicht einmal berücksichtigt, daß die Suche allein schon den Heuhaufen kräftig durcheinander wirbelt und kein Halm mehr auf seinem Platz ist. Genauso beim Screening: Was bedeutet es für eine Frau, die regelmäßig zur Mammographie geht, daß sie bereits bis zur zehnten Mammographie eine Chance von 49% hat, sich einen falsch-positiven Befund einzuhandeln, wie ein anderer Lancet-Artikel gezeigt hat (2)?

Aber nicht nur für die Organisatoren eines solchen Screening-Programms ist die Qualitätsanforderung verheerend. Dasselbe gilt für jeden, der in einer Studie die Qualität eines Screening-Programms überprüfen will. Die geringe Prävalenz hat zur Folge, daß begleitende Studien wesentlich größer sein müssen als therapeutische Studien zu ähnlichen Effektgrößen. Die Studien müssen randomisiert durchgeführt werden, da sonst der Selbst-Selektions-Bias nicht zu beherrschen ist (paradoxerweise gehen die Gesunden zumeist eher zur Screening-Untersuchung als die Kranken). Das bedeutet, daß man eine große Gruppe von Personen nicht in das Screening-Programm, sondern nur in das Auswertungsprogramm einbeziehen darf. Parallel zur Einführung eines flächendeckenden Programms dürfte das kaum möglich sein. Wegen der erforderlichen langen Nachbeobachtungszeiten (der Nutzen vieler Programme ist erst langfristig zu erwarten) müssen deshalb solche Studien früh aufgelegt werden. Als Ausweg wird deshalb gelegentlich vorgeschlagen, alle Personen zur Vorsorge einzuladen, aber nur eine zufällige Hälfte der Test-positiven zu therapieren. Dieser Vorschlag ist jedoch methodisch und ethisch unzureichend: die eventuell vielfältigen Nebenwirkungen des Screening-Programms bleiben außen vor, und es ist etwas anderes, einer gesunden Person ein zweifelhaftes Screening-Programm vorzuenthalten oder einer Test-positiven Person die Chance auf eine als wirksam erwiesene Therapie zu verweigern.

Ein fast unüberwindliches Problem ist zudem die eigentlich erforderliche Verblindung, die in der Praxis kaum umzusetzen ist: Über Jahre bis Jahrzehnte müßten alle Fährnisse, denen die Gescreenten unterliegen, ohne Kenntnis der Gruppenzugehörigkeit („blinded reading“) bewertet werden, auch solche, die nicht unmittelbar mit der betrachteten Krankheit, aber eventuell mit dem Screening und seinen Folgen zusammenhängen können, z.B. Biopsie- oder Bestrahlungsfolgen (in diesem Punkt hatten die Mammographie-Cochrane-Reviewer besondere Bedenken bei den bisher durchgeführten Studien). In der Bilanz ist die Bewertung eines Screening-Programms ein ähnlich schwieriges Vorhaben wie die Durchführung des Screening-Programms selbst. Hier stößt die wissenschaftliche Forschung an ihre Grenzen.

Wie also umgehen mit dem Anspruch der Evidenz-Basiertheit? Ihn als unrealistisch aufgeben? Keineswegs. Schauen wir uns die aktuellen Reviews zu therapeutischen Ansätzen an: Viele heute geklärte Fragen hätte man auch mit einem Evidenzansatz vor 10 Jahren gar nicht klären können, weil die Evidenzbasis schlicht unzureichend war. Trotzdem hat das Streben nach dieser Basis Erfolg gehabt. Wir sollten uns allerdings zugeben, daß es Bereiche gibt, in denen mit menschenmöglichen Anstrengungen eine befriedigende Evidenzlage nicht zu erreichen ist. Das Mammographie-Screening gehört möglicherweise dazu: Die Cochrane-Reviewer haben einen zur Bewertung der Folgen für die Gesamt-Mortalität erforderlichen Stichprobenumfang von 2 Millionen Frauen errechnet, der wohl

Diese Publikation ist urheberrechtlich geschützt. Vervielfältigung sowie Einspeicherung und Verarbeitung in elektronischen Systemen ist nur mit Genehmigung der Herausgeber des ARZNEIMITTELBRIEFS gestattet.

nicht mehr zu erreichen ist. Die Fortsetzung der bestehenden Studienprogramme wird uns nicht aus dem Dilemma befreien. Während diese Zeilen entstehen, wiederholen sich die Vorgänge aus dem Jahr 2001 auf der Basis weiterer Studienergebnisse für die Altersgruppe der 40-49jährigen: Die Ausgabe der Ann. Intern. Med. vom 3. September 2002 enthält nebeneinander die neuesten Analyse-Ergebnisse zur kanadischen Screening-Studie mit einer negativen Empfehlung für Frauen zwischen 40 und 49 Jahren (3, 4) und einen neuen Review der U.S. Preventive Services Task Force mit einer positiven Empfehlung zur Mammographie (5, 6) auf gleicher Datenbasis, unterschiedslos für Frauen zwischen 40 und 69 Jahren. Beide Seiten haben aufgrund der laufenden Debatte statistisch erheblich aufgerüstet, näher gekommen sind sie sich dadurch nicht. Gleich zwei Editorials können angesichts dieser Situation auch nur die Krise der Evidenz-basierten Medizin konstatieren (7, 8).

Heißt das, auf das Mammographie-Screening zu verzichten? Nicht unbedingt. Vielmehr ist hier in Anbetracht der unzureichenden Evidenzlage eine politische Entscheidung zu treffen: Fürchten wir eher die Gefahr, eine möglicherweise für viele Frauen hilfreiche Maßnahme zu unterlassen, oder das Risiko eines unnützen Aufwandes mit unangenehmen Nebenerscheinungen? Die Antwort wird von der Einschätzung der Evidenzlage und der persönlichen Risikobereitschaft abhängen. Solche Bündelung von unterschiedlichen Präferenzen ist eine primäre Aufgabe der Politik.

Leider ist jedoch Politik an dieser Stelle (wie es wohl auch sonst gelegentlich vorkommen soll) nicht ehrlich: Sie schiebt, wie wir lesen konnten, die angebliche Evidenzlage vor. Und hier kommen wir zu dem eigentlichen Dilemma, in dem wir uns mit der Evidenz-basierten Medizin befinden: Nicht ihre Begrenztheit ist das Problem, sondern ihre Verwendbarkeit. Es scheint auf den ersten Blick, als hätte die Evidenz-basierte Medizin die Politik durchdrungen. Doch das Gegenteil ist der Fall: Die Politik hat den Nutzen des Kampfbegriffs „EbM“ erkannt und beginnt, die Evidenz-basierte Medizin zu durchdringen. Die weltweite EbM-Bewegung hat leider auch selbst zu dieser Entwicklung beigetragen. In dem Bestreben, rationale Prinzipien in der medizinischen Praxis durchzusetzen, war eine all zu offene Diskussion der Grenzen dieses Ansatzes verständlicherweise nicht immer opportun. Nun muß verloren gegangenes Terrain wieder eingeholt werden. Die Stellungnahmen des Deutschen Netzwerkes zur Debatte um das Mammographie-Screening lassen erkennen, daß zumindest hier zu Lande die Herausforderung angenommen wird.

Aber läßt sich das Terrain noch einholen? Ein Blick auf die Homepage der Deutschen Gesellschaft für Senologie (www.senologie.org), die das Mammographie-Screening propagiert und deren Mitglieder wohl auch überwiegend davon profitieren würden, zeigt, wie weit wir gekommen sind. Die Evidenzlage bei der Mammographie wird hoch gelobt und die Notwendigkeit der Einführung eines flächendeckenden Programms daraus abgeleitet. Die mangelnde Evidenzlage bei der Brustkrebs-Selbstuntersuchung wird beklagt und deshalb der Nutzen in Zweifel gezogen, vor allem, wenn sie an die Stelle der Mammographie treten sollte. Trotzdem soll sie weiter empfohlen werden, da sie Frauen für das Thema Vorsorge sensibilisiert und mammographiebereit macht. Zur klinischen Untersuchung durch den Arzt lägen keine randomisierten Studien vor. Trotzdem wird diese Untersuchung dringend empfohlen, u.a. weil sie helfe, falsch negative Mammographie-Befunde zu erkennen. Die Untersuchung sei auch deshalb so nützlich, weil Ärzte auch Tumorgrößen von unter einem cm ertasten könnten, eine Fähigkeit, die Frauen bei der Selbstuntersuchung glatt abgesprochen wird.

Diese Publikation ist urheberrechtlich geschützt. Vervielfältigung sowie Einspeicherung und Verarbeitung in elektronischen Systemen ist nur mit Genehmigung der Herausgeber des ARZNEIMITTELBRIEFS gestattet.

Woher die Autoren das zu wissen meinen, bleibt unklar; in diesen Abschnitten kommt vorsichtshalber das Wort „Evidenz“ nicht vor. Hier wird mit dreierlei Maß gemessen. Wir sehen hier einen interessengeleiteten, willkürlichen Umgang mit den Anforderungen an die Evidenzlage einer Untersuchung, die den ansonsten inflationären Umgang mit dem Wort „Evidenz-basiert“ wirkungsvoll kompensiert.

Wir erkennen hier das Kardinalproblem der Evidenz-basierten Medizin, das sie seit Anbeginn begleitet: Sie trifft auf Menschen (Ärzte, Standesvertreter, Hersteller, Regulatoren, Patienten), die sich in vielen Punkten bereits positioniert haben. Die Unschuld des Nicht-Wissens und Nicht-bereits-gehandelt-Habens weisen überwiegend nur die Jungen und die Newcomer auf. Für uns alle anderen gilt die Frage: in welchem Umfang lassen wir uns auf neue Evidenz und neue Argumente ein? Nur wenn wir bereit sind, uns durch die Evidenz-basierte Medizin bewegen zu lassen, wird die Leitidee etwas bewirken. Anderenfalls verkommt sie zum Spielball der Interessen. Es liegt an uns, ob unsere Kinder beim Begriff „Evidenz-basierte Medizin“ mißtrauisch werden, weil sie einen politisch motivierten Durchsetzungstrick oder einen Verkaufsversuch vermuten, oder ob sie darin etwas sehen, woran man glauben kann.

Literatur

1. Olsen, O., und Gøtzsche, P.C.: Lancet 2001, **358**, 1340.
2. Elmore, J.G., et al.: N. Engl. J. Med. 1998, **338**, 1089.
3. Canadian Breast Cancer Screening Study-1: Ann. Intern. Med. 2002, **137**, I28.
4. Miller, A.B., et al. (Canadian Breast Cancer Screening Study-1): Ann. Intern. Med. 2002, **137**, 305.
5. U.S. Preventive Services Task Force: Ann. Intern. Med. 2002, **137**, I47.
6. Humphrey, L.L., et al.: Ann. Intern. Med. 2002, **137**, 347.
7. Goodman, S.N.: Ann. Intern. Med. 2002, **137**, 363.
8. Sox, H.: Ann. Intern. Med. 2002, **137**, 361.

Diese Publikation ist urheberrechtlich geschützt. Vervielfältigung sowie Einspeicherung und Verarbeitung in elektronischen Systemen ist nur mit Genehmigung der Herausgeber des ARZNEIMITTELBRIEFS gestattet.