

Evidence-based Medicine und zerhackte Rattenschwänze

Das Prinzip der externen Validität*

J. G. Schmidt

«Alles kann Gift oder Heilmittel sein, auf die Dosis kommt es an.» (Paracelsus)

«Alle Medizin nützt, wenn sie richtig eingesetzt wird und der beschränkte Geltungsbereich verstanden wird. Alle Medizin schadet, wenn sie falsch und ausserhalb ihres Geltungsbereiches eingesetzt wird.» (Informationsblatt, Stiftung Paracelsus Heute)

«Die Signatur ist ganz ausser Brauch gekommen und vollständig vergessen worden. Daraus ist grosses Irrsal gefolgt: Bisher hat kein Arzt (...) nach rechtem Grund beschrieben, was in den Dingen der Natur sei. Alle haben sie nur nach dem Hörensagen geschrieben.» (Paracelsus)

Für fast jeden Standpunkt lässt sich ein Paracelsus-Zitat finden aus jener Umbruchszeit vor 500 Jahren, die sich mit nicht mehr gültigen und noch nicht gültigen Wahrheiten zwischen alt und neu befand. Die in vielem sehr gleiche Entwicklung bedauern geradlinig strukturierte Menschen heute als postmoderne Beliebigkeit, die Wanderer zwischen den Welten sprechen ahnungsvoll von Vernetzung, und die meisten haben keine Ahnung, was passiert. Kein Wunder, dass sich in diesem bereits zur Theorie arrivierten Chaos die Medizin mehr anhand von Gerüchten denn anhand durchdachter Fakten entwickelt (vgl. Evidence-based Medicine und der Wandel von «Qualität», SÄZ Nr. 52/97), und wieder «alle nach dem Hörensagen» schreiben und lehren.

Die Externe Validität als querdenkerische Revolution

Wenn nun von zerhackten und ganzen Rattenschwänzen die Rede ist, wissen einige vielleicht bereits, dass es dabei um die chinesische Medizin geht oder auch um Uriella. Wenn Beat Gurtner sich über

* Nach einem Beitrag am BSV-Workshop «Kriterien zur Beurteilung des Nutzens von komplementärmedizinischen Methoden» im Dezember 1997

Korrespondenz:
Dr. Johannes G. Schmidt,
Allgemeinpraxis & Akupunktur / Chinesische Medizin,
Institut für Klinische Epidemiologie, Stiftung «Paracelsus Heute»,
Praxiszentrum Meinradsberg,
CH-8840 Einsiedeln

den «ausgebildeten Kranken» in einer durchaus lezenswerten Glosse in der Ärztezeitung äussert [1], den vernünftigen A-Patienten darstellt, der im Internet zu surfen weiss, und sich über den eher unvernünftigen B-Patienten (B wie Bio) amüsiert, der neben zerhackten Rattenschwänzen auch noch Nashornpulver und weiss der Teufel was ingestiert, so ist ab diesem Amusement die «Vernunft» aber allzu trügerisch und schnell verteilt.

Ich gestehe, ich habe auch schon Nashornpulver verschrieben und betrachte die heutigen Verhältnisse und Fragen der Rationalität als weit komplexer. Allzugerne wird übersehen, dass der wohl löbliche aber und dogmenorientierte Drang der offiziellen Medizin, «Scharlatanerie» auszugrenzen, just Scharlatanerie in einem Ausmass generiert hat, welches die Uriella-Medizinen in den Schatten stellt. Gerade weil die offizielle Medizin ihre Macht als orthodoxe Kirche langsam verliert, können wir aber heute den Mut haben, statt die doch eher unbedeutenden Uriellas einmal die viel einflussnehmenderen Halsteds, die Cholesterin-Guideliner (bzw. -hardliner) sowie die übrigen Labor-kosmetiker, die Hormonvernarrten, allerlei Check-upper, Früherkennungs-Zeloten und den ganzen Rattenschwanz in die Galerie der Scharlatane einzuordnen. Ein solches Gedankenspiel wirkt dann plötzlich nicht mehr so amüsiert und billig, obwohl das heute gültig werdende Evidenzverständnis diese mögliche Sicht dereinst zu mehr als einem Gedankenspiel werden lassen wird. Jedenfalls führt Evidence-based Medicine bzw. die Klinische Epidemiologie oft nicht einfach zu mehr (Internet-)Wissen, sondern vielmehr zur Offenbarung einer ungeahnten medizinischen Ignoranz, weil gültige und trugschlussfreie Evidenz auch für die allermeisten Interventionen der sogenannten Schulmedizin schlicht fehlt [2] und weil viele entscheidende, pragmatische Forschungsfragen – insbesondere die einer Nosologie von Placebowirkungen und einer Nosologie des Wirtes [3] – bisher nicht Eingang in die mit Untergangssängern besetzten Hochburgen der klinischen Forschung («publish or perish») gefunden haben (vgl. Evidence-based Medicine und «Placebo», SÄZ Nr. 42/97). Das Wissens-Chaos ist in der Tat perfekt. Und wenn der sprichwörtliche Flügelschlag eines chaostheoretischen Schmetterlings in Australien in einem ganz bestimmten Moment einen amerikanischen Tornado bewirken kann, wieso soll nicht auch ein von einem globalsichtigen Therapeuten verschriebener zerhackter Rattenschwanz in einer ganz bestimmten Patientensituation eine günstige Wirkung zeigen können?

Der friedensstiftende Engel für jene, welche dieses unerhörte Gedankenspiel in Aufruhr versetzen könnte, ist nun die wissenschaftlich schon immer entscheidend gewesene Frage der ganz bestimmten Situation, das Prinzip der externen Validität. Während etwa beim Routine-Ultraschall in der Schwangerschaft die interne Validität angeblicher Nutzenbeweise aufgrund von Störfaktoren und Biases (Surrogat-Bias, Confounding Bias, Selection Bias usw.) nicht gegeben ist, geht es bei der externen Validität um die begrenzte Gültigkeit von in sich gültigen

Studienergebnissen. Die Gültigkeit der Ergebnisse ist beschränkt auf die für eine Studie ausgewählte Patientengruppe mit ihren ganz bestimmten Krankheitsverhältnissen sowie auf die ganz bestimmten Studienbedingungen (erreichte Qualität der Intervention usw.). Die Cholesterinbestimmer und Laborwertverbesserer, die Hormonverschreiber und Blutdruck-Checker dürfen aufatmen: In ganz bestimmten Situationen können solche Interventionen nützlich sein. Dabei müssen insbesondere Vorbehandlungs-Risiko («base line risk») [4] oder die Vortest-Wahrscheinlichkeit («pre-test likelihood») [5] als entscheidende Kriterien der ganz bestimmten Patientensituation verstanden werden.

Während das ungezielte Cholesterin-Screening (mit den nachfolgenden Fibrat-Therapien) bisher mehr Menschen umgebracht als gerettet hat, weil die Mortalität bei den viel zahlreicheren Niedrigrisiko-Personen durch die Behandlung anstieg [6], so ist die Medizin inzwischen langsam dabei zu lernen, dass das Gesamtrisiko der ganzen Patientensituation entscheidet und dass es einen Cholesteringrenzwert und ein normales oder pathologisches Cholesterin gar nicht gibt [7, 8]. Während ein generelles Cholesterinscreening in die obengenannte Galerie gehört, ist der gezielte Einsatz von Statinen bei etablierter KHK oder vielen gehäuften Risikofaktoren recht nützlich, obwohl diese gezielte Cholesterinbehandlung heute paradoxerweise vernachlässigt wird. Die Berücksichtigung der externen Validität – was Wissenschaft erst als solche ausmacht! [9] – lässt uns am sehr gut dokumentierten Modellbeispiel Cholesterin heute erkennen, dass eine «wirksame» Therapie ausserhalb ihres Patienten-Geltungsbereiches umbringen kann, während sie bei gut berücksichtigtem Geltungsbereich Leben retten kann. Und die Behandlung eines «normalen» Cholesterins beim wirklichen KHK-Patienten wirkt lebensrettend, während die Behandlung eines «pathologischen» Cholesterins beim gesunden Wirt schaden kann. Zu Ende gedacht, sind «pathologische» Befunde bei geringer Vortest-Wahrscheinlichkeit oder bei geringem Vorbehandlungs-Risiko als biologisch normal zu werten und nicht als (falsches) Krankheitszeichen, das uns kraft unser unprofessionellen Sozialisation schon wieder zum unaufgeklärten und trivialen Exorzismus verführen will. Wieso können wir das Pathologische oder nur schon das vermeintlich Pathologische nicht (leben) lassen? Wieso erkennen wir nicht einmal den Wahrscheinlichkeits-Charakter und die Fehlerhaftigkeit von Diagnostik? Nicht das Resultat «pathologisch» entscheidet, sondern die Testsituation mit der Vortest-Wahrscheinlichkeit [5, 10]! Die moderne, aber gläubische Mystifizierung von apparativer «Abklärung» ist jedenfalls weit mehr abhängig von persönlicher Reife und Ängsten, dem Sterben (und dem «Pathologischen») ins Auge zu blicken, als von der wie auch immer «naturwissenschaftlichen» oder «alternativmedizinischen» Therapieschule, der man (angeblich) folgt.

Wenn auch die revolutionäre Tragweite der wiederentdeckten externen Validität vermutlich noch kaum verstanden wird, so ist dieses Prinzip heute ein

omnipräsenter Bestandteil von Evidence-based Medicine: Auf fast jeder Orientierungskarte im Handbuch von David Sackett und Konsorten [11] steht die Frage: «Can you apply this [internally] valid, important evidence in caring for your [particular] patient?»

Methodenpluralismus: ein Streit von gestern

Wie wir sehen, ist die Postmoderne nicht einfach beliebig, sondern sie verlangt dezidiert die Berücksichtigung der externen Validität, des genauen Geltungsbereichs. Natürlich gilt dies auch für Heiltheorien, seien diese nun «naturwissenschaftlich» oder «alternativmedizinisch». Die Klinische Epidemiologie, wie sie sich heute irreversibel durchgesetzt hat, anerkennt schon lange einen Methoden-Pluralismus in dem Sinne, dass nicht die Erklärbarkeit und (naturwissenschaftliche) Plausibilität einer Wirkung, sondern die sorgfältige Dokumentation der Existenz einer Wirkung entscheidend ist. Komplementärmedizinische Wirkmodelle sind von der heute gültigen Wissenschaft längst als denkbar akzeptiert, entschieden gefragt wird hingegen nach sorgfältigen Daten, welche die postulierte Wirkung auf den Patienten und den Nutzen dokumentieren. Und weil die pragmatische Klinische Epidemiologie auch immer entschiedener nach der Praxisrelevanz nosologischer Einteilungen fragt, und weil diese «Nosologie» etwa in Form der Baseline-Charakteristika in klinischen Studien auch Wirtsmerkmale erfasst, ist auch die Integration verschiedenster Medizinrichtungen methodisch bereits vorgegeben [9]. Mit der Klinimetrie ist auch das vom Patienten erlebte Erleiden (statt seiner Laborwerte) Massstab für Behandlungserfolg geworden. Sogar die orthopädischen Chirurgen der Schweiz bereiten heute eine Abkehr von herkömmlichen biomechanischen und anatomischen Erfolgskriterien hin zu Lebensqualitätsparametern wie der Alltagsfähigkeit vor.

Die Frage ist deshalb längst nicht mehr, wie angeblich linear, kausal-analytisch, verdünnt, unverdünnt, eingebildet oder autistisch Medizin ist, sondern wie wir mit dem bereits etablierten Methoden-Pluralismus überzeugend umgehen. Ein grosses Problem dabei ist, dass das dazu nötige, differenzierte methodische Wissen hierzulande noch viel zu dünn gesät ist, wodurch souveräne Visionen und Einsichten fehlen, welche die vergeudete Energie unkundiger und überholter Streitereien vermeiden könnten.

Wissenschaft kennt keine Grenze zwischen Schulmedizin und Komplementärmedizin

Der heute gültige wissenschaftliche Rigor und die Hinterfragung der Evidenz nach den Regeln der klinischen Epidemiologie haben auch für die Komplementärmedizin Gültigkeit. Dabei kann es unterschiedliche Schwerpunkte geben, die aber dennoch alle Medizin betreffen. Wenn etwa präventive Massnahmen jeglicher Provenienz in das Leben Gesunder eingreifen wollen, ist eine rigorose Evidenz unab-

dingbar, die beweist, dass diese Massnahmen wirklich mehr nützen als schaden. Dies geht in der Regel nicht ohne (aufwendige) randomisiert kontrollierte (Langzeit-)Studien. So muss bei der anstehenden Leistungsbewertung auf jeden Fall durchgesetzt werden, dass etwa biologische Testverfahren der Komplementärmedizin, die zum Beispiel zum Screening von Krebsdispositionen eingesetzt werden und Patienten die höchst problematische und einschneidende Etikette einer Krebsgefährdung verpassen, genauso sorgfältig evaluiert werden, wie wir das heute von aller Medizin verlangen müssen. Wenn aber «Das autistisch-undisziplinierte Denken in der Medizin» von Bleuler als offenbar zeitloser Knüller herhalten muss, um jegliche kritische Offenheit der Komplementärmedizin gegenüber zu diskreditieren und um den verbreiteten Irrtum einer Gleichsetzung der randomisiert kontrollierten Studie als Standard des Nutzensnachweises mit der Doppelblindstudie zu pflegen [12], dann muss daran erinnert werden, dass gerade die (angeblich) naturwissenschaftliche Medizin fast ihr ganzes präventivmedizinisches Armamentarium ohne adäquate Nutzensüberprüfung auf die breiten Massen losgelassen hat. Dies gilt auch für einen grossen Teil der zahlreichen tertiärpräventiven Behandlungen, die den medizinischen Alltag beherrschen und in vielen Bereichen eine vermutlich schädliche Medizin ausmachen.

Dass der Standard des Wirksamkeitsnachweises die Doppelblind-Studie sei, wird unglücklicherweise auch im Kriterienpapier der Komplementärmediziner [13] behauptet, obwohl dies schlicht falsch ist [14] und die Diskussion nur in eine unnötige Sackgasse führt. Ausser einigen ewig-gestrigen Homöopathen-Fressern und methodischen Halbwissern bestreitet heute niemand, dass die Doppelblind-Studie als besondere Form der randomisiert kontrollierten Studie für viele klinische Forschungsfragen ungeeignet ist. Die Doppelblindstudie ist letztlich nur ein Instrument zur Prüfung der internen Validität einer pharmakologischen Wirkungstheorie, für die Prüfung des Patientennutzens sind vielmehr sogenannte pragmatische randomisierte Studien mit praxisgerecht individualisierten Therapien erforderlich [9]. Offenbar dient das Festhalten an einem überholten methodischen Halbwissen auf beiden Seiten dazu, die andere Seite zu diskreditieren, um die liebgewonne Identifikation mit der alleinigen Wahrheit seiner Schule nicht verlieren zu müssen. Die Zeit des ideologischen Streits um Modell und Gegenmodell ist aber vorbei, die heutige klinische Forschungsmethodik will die begrenzte Gültigkeit (externe Validität) von Modell und Gegenmodell erhellen.

Deshalb gehört die sogenannte «wissenschafts- und medizinthoretische Grundposition der Komplementärmedizin», welche zur Kompensation anderer Lücken in das BSV-Dokument [13] Eingang gefunden hat, ins Museum. Eine Heiltheorie und Therapieschule ist nie an sich richtig oder falsch, wissenschaftlich dokumentierbar sind nur deren indikationsbezogene Leistungen im Einzelnen. Auch mit der Komplementärmedizin wird sehr häufig medikalisiert und ein Hilfsbedarf in Patienten projiziert, der mehr mit ei-

gener Anerkennungs-Bedürftigkeit als mit des Patienten Problem zu tun hat. «Grundpositionen» machen noch lange keine bessere Medizin und gekonntere Praxis aus. Wenn in der komplementärmedizinischen Praxis eine Therapiesituation «einmalig» sein soll [13], dann ist dies in einer guten Schulmedizin genauso der Fall! Wie will man denn mit solchen «Kriterien» erkennen können, ob eine Therapie einmalig nützlich oder einmalig schädlich oder überflüssig ist? Es ist sicher nicht falsch, solche Grundpositionen zu verdeutlichen; aber es ist falsch, sich damit von (einem Zerrbild) der Schulmedizin abgrenzen zu wollen. Es sind nun einmal «Schulmediziner» wie Feinstein und die McMaster-Fakultät, welche die methodischen Grundlagen der Klinischen Epidemiologie und der Klinimetrie entwickelt und mit langem Atem durchgesetzt haben, während aus der Komplementärmedizin nur Gegenmodell (bis zum geht nicht mehr), aber kaum integrative Methodik gekommen ist. Die Komplementärmedizin muss *dieser* Schulmedizin dankbar sein und sie nutzen statt noch länger ihre Wunden der Besonderheit zu pflegen. Die entscheidende Aufgabe wäre, wirklich zu formulieren, wie die an sich legitimen Grundpositionen im Kontext klinischer Forschung (mit den vielen Kompromiss-Zwängen) operationalisiert werden. Man wird dann erkennen, dass Komplementärmedizin und schulmedizinische Allgemeinmedizin mit den genau gleichen methodischen Problemen in der Planung praktisch relevanter klinischer Studien konfrontiert sind.

Randomisierte Studien, Fallstudien und klinisches Ermessen

Ohne Zweifel ist die wissenschaftliche Überprüfung komplementärmedizinischer Wirkungen bei intelligentem Design auch in randomisierten, vergleichenden Studien möglich, und dies ist schliesslich die einzige Evidenz, die gut vor einer Reihe von Biases geschützt ist. Zurecht spricht die Klinische Epidemiologie von einer Hierarchie der Evidenz. Aus den heute gut etablierten Kenntnissen der möglichen Fehlerquellen unkontrollierter Studien bleibt diese Hierarchie der Evidenz auch für die Komplementärmedizin prinzipiell gültig. Sicher sind primär sorgfältige Fallstudien sehr wichtig, schon um möglichst genaue Hypothesen zu formulieren. Die meisten Fallstudien in der Komplementärmedizin (wie auch in der Schulmedizin) werden jedoch methodisch schlecht durchgeführt, die Erfolgsbeurteilung ist voreingenommen, die Patientenselektion ist unklar und die Interpretation unsorgfältig. Wenn Fallstudien sorgfältig durchgeführt werden, sind sie für Verbesserungen von Hypothesen und für die Entwicklung neuer therapeutischer Erkenntnisse und deren Differenzierungen unverzichtbar. Ganze Fallserien oder gute Kohortenstudien können oft noch härtere Hypothesen liefern, die aber unter Umständen auch erst in kontrollierteren Studien weiter validiert werden können.

Und zweifellos können gute unkontrollierte Studien unter Umständen Hypothesen so weit eingren-

zen, dass bereits praktisch gültige Arbeitshypothesen entstehen. Aber es bleiben Arbeitshypothesen mit ihren inhärenten Unsicherheiten, was mit keiner (zum Teil sehr berechtigten) Kritik an der randomisierten Studie zugedeckt werden kann. Die Ergebnisse (praktisch relevanter) randomisiert kontrollierter Studien sind nicht nur in dem Sinn valider als sie Selektions- und andere Biases ausschalten, sondern auch indem sich die therapeutische Intervention frei von solchen Störfaktoren in ihrer unverfälschten Wirkung und Wirkgrösse zeigt, was für die praktisch wichtige Einschätzung der «Number needed to treat» unabdingbar ist. Ergebnisse randomisierter Studien sind aber Durchschnittsergebnisse, die im unklaren lassen, ob der Patient, den wir vor uns haben, wirklich von dieser Therapie profitiert. Aber diese Ergebnisse sind, wenn die Studien praxisgerecht angelegt sind, das Beste, was wir wissen, ausser wir haben im Einzelfall gut begründete Arbeitshypothesen (die kundig und nicht einfach durch Unsorgfältigkeit und Denkfehler zusammengebastelt worden sind), die einen anderen Entscheid nahelegen (eine gesicherte Therapiewirkung kann insbesondere auch durch eine unvernünftig grosse «Number needed to treat» relativiert werden, und eine gute private Hypothese kann dann legitimerweise die Oberhand gewinnen). Im Vergleich zu präventiven Versprechen sind Therapien zur Linderung eines Leidensdrucks oft durch ihren Erfolg bereits zu einem gewissen Grad gerechtfertigt. Besonders bei chronisch stabilen Störungen kann der Patient unter Umständen als seine eigene Kontrolle dienen, was dem Erkenntniswert einer randomisierten Kontrolle möglicherweise sogar überlegen ist. Die begrenzte Gültigkeit, aber auch Uersetzbarkeit der randomisiert kontrollierten Studie ist längst ein Thema an gut informierten Biometrik-Symposien [15].

Im Patienten-Einzelfall entscheidet am Schluss immer das klinische Ermessen, und nicht randomisierte Studien alleine. Aber dies ist kein Argument, welches kontrollierte Studien in Frage stellen könnte. Das Problem mit dem klinischen Ermessen ist, dass dieses möglichst explizit geschehen sollte, und dass dieses klinische Ermessen auf sorgfältigen Beobachtungen und möglichst validen Forschungsergebnissen und auf einer logisch-kritischen Hinterfragung der Evidenz beruhen muss. Auch «intuitives» Entscheiden und Therapieren wird damit nicht grundsätzlich in Frage gestellt, die Berufung auf «Intuition» befreit aber nicht vor sorgfältiger und unvoreingenommener (d.h. unabhängiger oder verblindeter) Beurteilung der Ergebnisse. Es darf nicht übersehen werden, dass alles, was ausbildnerisch weitervermittelt werden kann, dadurch auch im klinimetrischen Sinne messbar und in der klinischen Forschung überprüfbar wird. Die vor allem von Feinstein entwickelte Klinimetrie befasst sich mit der kritischen und systematischen Bemessung des klinischen Erfolgs, man könnte auch sagen mit dem Ausmass des Leidens und der Leidensverminderung. Wenn dabei das subjektive Patientenurteil oft wichtiger ist als die heute noch im Vordergrund stehenden apparativen Daten, so enthält dies durchaus auch objektivierbare Information,

wenn etwa die Absentismus-Tage oder die schlaflosen Nächte gezählt werden.

Es gibt nur gültige Fragen, aber keine endgültigen Studien

Wie auch Feinstein und Horwitz eben wieder einmal fundiert dargelegt haben [16], wäre ein Abstützen der medizinischen Praxis auf bisher vorliegende Evidenz aus randomisiert kontrollierten Studien falsch und grob einseitig, weil Studienergebnisse nur Durchschnittsergebnisse sind, welche wichtige klinische Gesichtspunkte und Differenzierungen übergehen, und weil aufgrund einer wissenschaftlich obsoleten Labormessgläubigkeit oft auch nicht die klinisch entscheidenden Fragen beantwortet sind. Für die Frage der Leistungsbeurteilung aller Medizin bedarf es deshalb vor allem eines Katalogs kritischer Fragen, welche die Qualität und Gültigkeit vorgebrachter Studien und Erfahrungen aufgrund ihrer jeweilig möglichen Störfaktoren hinterfragen. Und letztlich steht nicht der Studientyp an sich zur Diskussion, sondern das klinische Problem, und aufgrund der genauen klinischen Frage ergibt sich erst das geeignete Studien-Design bzw. die geeignete Kaskade konsekutiver Studien. Wie gesagt, gilt es zum Beispiel zu unterscheiden, ob eine Massnahme einen bestehenden Leidensdruck mildert, oder ob man eine Kondition ohne Leidensdruck mit einer präventiven Absicht angehen will. Für den Praktiker aller Medizinen sind gute Fallstudien häufig wertvoller als randomisierte Studien, um seinen Patienten zu erkennen und bei ihm eine Lösung für dessen Schmerzen oder Krankheitsäusserungen zu finden. Vor allem bei chronisch stabilen Krankheiten zählt, was dem Patienten schliesslich eine Besserung gebracht hat. Bei präventiven Versprechen hingegen müssen Massnahmen sorgfältig in kontrollierten Studien überprüft werden, denn oft ist nicht nur die Frage entscheidend, ob überhaupt eine präventive Wirksamkeit da ist, sondern ob «wirksame» Interventionen insgesamt mehr nützen als sie iatrogene Probleme schaffen (vgl. Cholesterin-Screening). Wirksamkeit (efficacy) darf nicht schon mit Nutzen (effectiveness) gleichgesetzt werden.

In diesem Spiel gibt es alle Schattierungen klinischer Fragestellungen, die jeweils eine eigene, problemabhängig angemessene Strategie der Evidenzbeurteilung erfordern. Praxiskontext und Ebene der klinischen Frage entscheiden über die Angemessenheit verschiedener Studien-Arten. Und dabei kommt es nicht darauf an, ob die Frage aus der sogenannten Schulmedizin oder aus der Komplementärmedizin stammt. Man kann allenfalls feststellen, dass die typische Komplementärmedizin sich direkt am vom Patienten erlebten Leiden orientiert und damit viel weniger häufig eine Hinterfragung des Surrogat-Trugschlusses braucht. In diesem beschränkten Sinne ist Komplementärmedizin nach heutigem Evidenz-Verständnis in der Tat oft valider als die herkömmliche schulmedizinische Surrogat-Medizin, die ihre Erfolge an paraklinischen Testresultaten zu messen pflegt.

Vorläufige Beurteilungskriterien

Brauchbare Beurteilungskriterien für die Wirksamkeit komplementärmedizinischer Therapien könnten mindestens vorläufig aus einem Hinterfragungs-Katalog bestehen, mit welchem «Erfahrungen» und vorliegende Studien nach Störfaktoren, Biases und anderen möglichen Unzulänglichkeiten abgeklappert werden. Es braucht dann immer noch einen Ermessens-Entscheid über das Vorliegen einer (vorläufig) akzeptierbaren Nutzens-Evidenz. Auch für Interventionen aus der «Schulmedizin» wird ja so vorgegangen; es fordert ja auch niemand, dass der Anteil von über 90% schulmedizinischer Leistungen, der nicht durch gültige Evidenz untermauert ist, von der Kassenleistungspflicht gestrichen wird. Für die Bezeichnung von komplementärmedizinischen Therapien mit akzeptabler Nutzens-Vermutung kann vorläufig am besten eine methodisch geschulte, einer pluralistischen Sicht verpflichtete und praktisch erfahrene Kommission eingesetzt werden. Mit zunehmenden Forschungsergebnissen und zunehmenden methodischen Erkenntnissen sowie den Erfahrungen dieser Kommission könnten die Ermessens-Kriterien in einem transparenten Prozess laufend angepasst und präzisiert werden.

Für die Qualität und Effizienz der geleisteten Medizin weit entscheidender als die angeblichen oder echten «Qualifikationen» der schulmedizinischen oder komplementärmedizinischen Akteure ist im übrigen ein pragmatisch-intelligentes Zusammenspielen der verschiedenen Behandlungsmöglichkeiten. Dabei kann gleichzeitig auf Diagnostik und sogenannte «Abklärungen» verzichtet werden, die auf die Behandlungswahl und den Behandlungserfolg gar keinen entscheidenden Einfluss haben. Dazu gehört natürlich ein profundes Verständnis von *Medical Decision Making* [17] und des blossen Wahrscheinlichkeits-Charakters von Abklärungs-Ergebnissen und ein entsprechendes Entscheiden und Therapieren aufgrund von «Arbeits-Diagnosen», welche zudem noch die Teilnahmemöglichkeit des Patienten mitberücksichtigen.

Literatur

- 1 Gurtner B. Der ausgebildete Kranke. Schweiz Ärztezeitung 1998;79(15):638.
- 2 Field MJ, Lohr NN, eds. Institute of Medicine Guidelines for Clinical Practice: From Development to Use. Washington: National Academy Press;1992.
- 3 Feinstein AR. Krankheitserscheinungen am intakten Menschen: Was für eine Nosologie der Krankheitseinteilungen brauchen wir? 2. Wissenschaftliches Einsiedler Symposium. Einsiedeln: Stiftung Paracelsus Heute;1995.
- 4 Steurer J, Bucher HC, Schmidt JG. Evidence Based Medicine: Kritische Beurteilung einer «Therapiestudie». Praxis 1997; 86:1603-7.
- 5 Bucher HC, Schmidt JG, Steurer J. Evidence Based Medicine: Kritische Beurteilung einer Arbeit zu einem diagnostischen Test. Praxis 1998 [im Druck].
- 6 Davey Smith G, Song F, Sheldon TA. Cholesterol lowering and mortality: The importance of considering initial level of risk. Br Med J 1993;306:1367-73.
- 7 Schmidt JG. Cholesterinscreening: Die Irrationalität von Grenzwerten und die Berücksichtigung des Gesamtrisikos für eine rationale Therapie. In: Kochen MM. Rationale Pharmakotherapie in der Allgemeinpraxis. Berlin Heidelberg: Springer-Verlag; 1991. S. 67-81.
- 8 Sheffield table for primary prevention of CHD. Lancet 1996; 348:1352.
- 9 Schmidt JG. Die Vision einer pragmatischen klinischen Forschung oder das Ende der Diskussion über «Placebo» und «spezifische Wirkungen». 3. Wissenschaftliches Einsiedler-Symposium. Forsch Komplementärmed 1998;5(Suppl. 1): im Druck.
- 10 Schmidt JG. Leserbrief. Schweiz Med Wochenschr 1998; 128:599-600.
- 11 Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine - How to practice & teach EBM. New York/Edinburgh: Churchill Livingstone; 1997.
- 12 Geiser M. Leserbrief. Schweiz Ärztezeitung 1997; 78(50):1893-5.
- 13 Heusser P. Prüfkriterien zur Beurteilung des Nutzens von komplementärmedizinischen Methoden. BSV 30.10.1997.
- 14 Rothman KJ, Michels KB. The continuing unethical use of placebo controls. N Engl J Med 1994;331:394-8.
- 15 Abel U, Koch A. (eds.) Nonrandomized comparative clinical studies. Düsseldorf: Symposium Publishing; 1998.
- 16 Feinstein AR, Horwitz RI. Problems in the «Evidence» of «Evidence-based Medicine». Am J Med 1997;103:529-35.
- 17 Sox HC, Blatt MA, Higgins MC, Marton KI. Medical decision making. Boston: Butterworth-Heinemann;1998.