

J. G. Schmidt

Allgemeinpraxis, Institut für Klinische
Epidemiologie, Stiftung «Paracelsus Heute»,
Einsiedeln, Schweiz

Bestimmung von Stellenwert und Wirksamkeit komplementär- medizinischer Behandlungen: Brauchen wir eine «alternative» Forschungsmethodik?¹

Schlüsselwörter

Konventionelle Medizin
Naturwissenschafts-Credo
Trugschlüsse
Nosologie

Key Words

Conventional medicine
Science creed
Fallacies
Nosology

Zusammenfassung

Vor Paracelsus war das Disputieren in Latein als Erkenntnismethode zu einer Gesinnung verkommen und musste ersetzt werden. Das «Latein-Credo» von heute ist das Naturwissenschafts-Credo; überlegtes Sammeln von naturwissenschaftlichen Daten wird mit Naturwissenschaftlichkeit verwechselt. Der klinische Epidemiologe Feinstein verlangt erneut eine Wissenschaft vom Patienten und bemängelt, dass sich das «klinische Material» (der Patient) durch die einseitig pathologisch-anatomisch ausgerichtete Nosologie nicht treffend charakterisieren lässt. Zunehmend werden dramatisch anmutende Fehler «naturwissenschaftlicher» Diagnostik und Therapieziel-Festsetzung von innovativen Köpfen innerhalb der Schulmedizin erkannt und beschrieben, während bei Forschungsanstrengungen von Komplementärmedizinern diese Fehler oft noch übersehen werden. Es braucht keine grundsätzlich neue Forschungsmethodik für die Komplementärmedizin; es gilt vielmehr, die tiefgreifenden methodischen Entwicklungen der letzten Jahre in der konventionellen Medizin zur Kenntnis zu nehmen und damit kreativ umzugehen.

Summary

Determination of Status and Efficacy of Complementary Medicinal Treatments: Do We Need an 'Alternative' Research Methodology?

Before the time of Paracelsus, disputing in Latin language as an epistemological method for gaining knowledge had decayed to an attitude and thus had to be replaced. Today's 'Latin creed' is science creed, and inconsiderate collection of scientific facts is mistaken for scientific approach. The clinical epidemiologist Feinstein revives the call for a patient-adjusted science and faults that the 'clinical material' (i.e. the patient) cannot be characterized appropriately by an ill-balanced, pathologic-anatomically oriented nosology. Increasingly, innovatively thinking persons in the field of conventional medicine identify and describe dramatically appearing shortcomings of 'scientific' diagnostics and of the establishment of therapy objectives, whereas these shortcomings frequently are overlooked by complementary medicine researchers. We do not need a basically new research methodology for complementary medicine; rather it is important to take notice of the profound methodological developments in conventional medicine in recent years and to creatively get on with them.

¹ Vortrag im Rahmen der Schaffung der «Kolle-gialen Instanz für Komplementärmedizin (KIKOM)» an der Universität Bern, 31. August 1994.

1579 fand im Pariser «parlement» eine Verhandlung statt, angeklagt war ein Paracelsischer Drogist, der sich als Arzt ausgegeben und Patienten behandelt hatte. Die Klägerin war die Medizinische Fakultät. In der Verhandlung wurde rasch festgestellt, dass der Angeklagte kaum Latein beherrschte, und das Gericht kam nach den Beratungen zum Schluss, dass dieser ohne Kenntnisse der lateinischen Grammatik unmöglich etwas von Medizin verstehen könne. Dem Angeklagten *Roch de Baillif*, immerhin «médecin ordinaire» des Königs Henry III., half der Hinweis dabei nichts, auch *Hippokrates* oder *Avicenna* hätten in ihrer Muttersprache gelehrt. Dies berichtet uns *Pamela Smith* im «Bulletin of History of Medicine» [1].

Wie Sie natürlich erkennen, stand hinter dieser Auseinandersetzung nicht nur die Suche nach wissenschaftlicher Wahrheit, sondern es spielte sich ein Glaubenskrieg ab zwischen dem Establishment mit seinen Regeln und einem praktisch erfolgreichen Autodidakten, verflochten mit einem Machtkampf zwischen König und Parlament.

«Credoisierung» von Wissenschaft

Die Medizinische Fakultät von Paris war eine Bastion der Galenischen Humoralpathologie und anerkannte das Studium lateinischer Texte als einzige Forschungs- bzw. Erkenntnismethodik. Es wäre allerdings allzu simpel, die damalige Medizinische Fakultät einfach als unaufgeklärt, rückständig oder dumm ansehen zu wollen. Die Disputation kann eine durchaus nützliche Erkenntnismethode sein. Wir können uns zudem vorstellen, dass die Konvention des Lateins bzw. die Struktur der lateinischen Grammatik die spielerische Seite dieser Methode unterstützte und die Wirksamkeit der Methode im Gewinnen von Einsichten erhöhte. Wir können uns auch vorstellen, dass damals viele obskure Anschauungen autodidaktischer und autistischer Ärzte existierten, die in der Folge vollständig verschwunden sind. Wieso hätte die Fakultät auf Altbewährtes vorschnell verzichten sollen? In der weiteren Geschichte konnte sich die Paracelsisten-Schule allerdings halbwegs durchsetzen.

Lassen Sie mich den Versuch machen, die wenigen Elemente zu nennen, die den Erfolg der Paracelsischen Methode begründet haben könnten:

1. Paracelsus gab der Beobachtung das Primat über die praxisfremde Theorie und schaffte damit die Grundlagen für eine neue wirksamere Medizin. Man könnte sagen: Er führte die Methode der kontrollierten Beobachtung ein.
2. Weil Paracelsus erkannte, dass beobachtbare Wirkungen für den Patienten entscheidender sind als die Stringenz und Konformität von Wirkungstheorien, konnte er eine zwanglose Integration unkonventioneller Heilverfahren in seine ärztliche Praxis vornehmen.
3. Paracelsus durchschaute als bäuerisch geprägter Landmensch von Einsiedlern die Ängste, die sich hinter der wortgewandten lateinischen Gelehrtheit an den damaligen

städtischen Medizinfakultäten verbarg, und er erkannte dadurch den Grund irrationaler Medizin.

Im Paris des ausgehenden 16. Jahrhunderts war das Latein zur Gesinnung und zu einem Kastenmerkmal verkommen und seine ehemals schöpferische Funktion ging damit verloren. Die «lateinische Methode» war nurmehr äusserliches Merkmal ohne Bezug zum praktischen Gegenstand, sie *musste* ersetzt werden. Ich möchte den Versuch wagen, das «Latein-Credo» von heute beim Namen zu nennen: Es ist das Naturwissenschafts-Credo. Die Methode der Naturwissenschaft, die damals das als dunkel bezeichnete Mittelalter abgelöst hat und in der Tat ein fruchtbares neues wissenschaftliches Zeitalter begründet hatte, scheint heute vom selben Schicksal der «Credoisierung» ereilt zu werden. Lassen Sie mich mit einer Andeutung diese Aussage untermauern: Die gegenwärtige Forschungspraxis besteht häufig gar nicht in Naturwissenschaft, sondern im Sammeln von naturwissenschaftlichen Daten. Dies wird aber von unbedachten Schülern dieses Systems gerne mit Naturwissenschaftlichkeit verwechselt.

Die moderne klinische Epidemiologie

«Der klinische Epidemiologe Feinstein bemängelt die Annahme, dass sich die intellektuellen Herausforderungen der medizinischen Wissenschaft ausschliesslich im Labor abspielen würden und die Aufgabe des nichtforschenden Arztes lediglich darin bestünde, die Entdeckungen der Laborforschung in die «Kunst» der Patientenbetreuung zu integrieren. Feinstein schlägt zur Ergänzung der medizinischen Grundlagenforschung eine Wissenschaft vom Patienten vor». Dies schreibt Michael Kochen in seinem Allgemeinmedizin-Lehrbuch [2]. Und weiter: «So wird der Begriff «hart» üblicherweise auf Informationen angewendet, die unwiderruflich sind (z. B. der Tod), von archivierbaren und wiederholbaren Proben stammen (Blut, Zellen, Gewebe, Röntgenaufnahme), auf objektivierbaren Beobachtungen (meist durch eine Maschine) beruhen und in quantitativen Dimensionen ausgedrückt werden können (z. B. Blutzucker = 112 mg/dl).» Kochen ist konventioneller Arzt, der keine Komplementärmedizin praktiziert.

Feinstein hat auch auf dem Einsiedler Symposium [3] am Beispiel solider Karzinome die Punkte aufgezählt, die bei einer konventionellen Krankheitsdiagnose «verloren gehen», obwohl sie für praktische Belange sehr wichtig sind und durch alle möglichen zusätzlichen Daten wie Histologie, Rezeptorstatus, Laborparameter etc. nicht ersetzt werden können. Es sind:

1. Die *Art der Beschwerden* (pattern of symptoms); so stellt ein Patient mit einem asymptomatischen Karzinom einen ganz anderen Fall dar als ein Patient mit dem gleichen Karzinom und dem gleichen Stadium, der sich bereits krank fühlt.
2. Die *Ausprägung der Symptome* (severity of symptoms); ein Patient, der viel Gewicht verloren hat, stellt einen anderen Fall dar als einer ohne Gewichtsverlust.

3. Die *Wachstumsgeschwindigkeit* (chronometry of events or auxometry); eine Frau mit dem genau gleichen Stadium-II-Mammakarzinom, welche ihren Tumor seit 20 Jahren hat, unterscheidet sich gewaltig von einer Frau, welche vor einem Jahr noch einen unauffälligen Palpationsbefund aufwies.
4. Die *Ko-Morbidität*; 2 Patienten mit dem genau gleichen Karzinom unterscheiden sich wesentlich, wenn der eine zusätzlich etwa eine dekompensierte Leberzirrhose aufweist.
5. Der *funktionelle Status* (functional capacity) eines Patienten; die individuelle Anpassungsleistung an einen Krankheitszustand in der Verrichtung täglicher Abläufe ist oft viel entscheidender für den Verlauf als alle klassischen klinischen Parameter und Zusatzdaten aus Labor und anderen Verfahren.

Das soziale Netz etwa bestimmt die Prognose eines Mammakarzinoms in höherem Ausmass als alle therapeutischen und Früherkennungsmöglichkeiten. So weisen 2 Untersuchungen [4, 5] auf eine markante prognostische Bedeutung des sozialen und emotionalen Netzes auf das Überleben von Brustkrebspatientinnen bzw. von schweren Lebensereignissen auf die Rezidivrate hin (relatives Risiko bei schweren Lebensereignissen etwa 5); und erste Daten aus einer kontrollierten Studie zeigen eine Verdoppelung der Überlebenszeit unter «supportive group therapy» bei Patientinnen mit metastasierendem Mammakarzinom [6]. Demgegenüber zeigt sich unter der Früherkennung ein relatives Mortalitätsrisiko von 0,8 [7]; zytostatische Therapien zeigen eine Wirkung ähnlich geringen Ausmasses. Leonard Syme hat die überragende Rolle der individuellen Schicksalsbewältigung (control of destiny) als prognostischen Faktor bei fast allen von ihm untersuchten Krankheiten dokumentiert [8].

Feinstein verlangt und entwirft eine Wissenschaft vom Patienten [9] – man könnte durchaus auch sagen eine Naturwissenschaft vom Patienten, wenn dieser Begriff heute nicht durch eine Gesinnung okkupiert und verfälscht wäre. Auch Feinstein ist konventioneller Arzt, Internist und Statistiker. Seine Kritik rührt nicht von einer alternativmedizinischen Praxis her, sondern vielmehr daher, dass er als sorgfältiger Forscher, Kliniker und Statistiker die Fragilität vieler statistischer Rituale durchschaut (es versteht sich, dass er daraus nicht eine diffuse Anti-Statistik-Haltung ableitet).

Klassische Diagnosen und die Praxis

Woraus besteht in der Praxis das «klinische Material» (um es etwas salopp zu formulieren)? Es sind einzelne Individuen, für die vom Arzt ein Behandlungsweg bzw. Heilungsweg aufgefunden werden muss. In kontrollierten Studien wird – neben prognostisch bedeutsamen Baseline-Charakteristika – wenn möglich immer auch die nichtspezifische Mortalität überprüft, um Gewähr für eine Gruppenvergleichbarkeit zu haben. So können insbesondere Studien mit einer historischen Kontrolle

aus unvergleichbaren Gruppen bestehen, und die Überprüfung der Gesamt mortalität – der «Gesamtsterbeprogno» – ist deshalb ein entscheidendes Kriterium der Gruppenvergleichbarkeit. So wird beispielsweise in einem Schweizerischen Fachblättlein eine angeblich nachgewiesene Reduktion der Brustkrebs mortalität durch die Brust-Selbstuntersuchung um 29% im Vergleich zu einer historischen Kontrollgruppe behauptet [10]; die Publikation der gleichen Studie in einer internationalen Zeitschrift mit adäquater methodischer Betreuung geht dann aber auch auf den praktisch gleichen Unterschied in der Gesamt mortalität ein, welcher zeigt, dass Studiengruppe und Kontrollgruppe nicht vergleichbar waren. In dieser Fachzeitschrift heisst es dann: «... selection bias, inherent in any observational study of screening, provided an alternative explanation for the findings» [11]. Die konventionelle Nosologie, welche sich vorwiegend an morphologische Einteilungskriterien anlehnt, kennzeichnet in keiner Weise prognostisch einheitliche Krankheitsfälle, d.h. Krankheiten mit identischer Natur oder identischer «natural history». Die heutige Nosologie ist eine Konvention und entspricht als solche einer willkürlichen Konstruktion, die nicht auf ewig eine Allgemeingültigkeit beanspruchen wird.

Die Einheitlichkeit des klinischen Materials kann möglicherweise besser durch eine einheitliche *Prognose* als durch eine einheitliche *Diagnose* gekennzeichnet werden. Es entstünde eine neue und sehr verschiedene Nosologie, wenn wir eine Krankheitseinteilung anhand praktisch-klinischer Kriterien vornehmen wollten. Das «klinische Material» in einer Allgemeinpraxis – fern vom idealtypischen und oft idealisierten «Lehrbuchmaterial» der grossen Kliniken – lässt sich kaum je durch eine konventionelle Diagnose treffend charakterisieren. In der Praxis treffen wir auf – wie bereits gesagt – ein einzelnes Individuum, für welches ein Behandlungsweg bzw. Heilungsweg gefunden werden muss.

Jede klinische Behandlung entspricht in der Tat jedesmal einem Experiment, wobei Arbeitshypothesen gebildet, wieder revidiert oder verworfen werden. Das entscheidende Kontrollkriterium ist dabei eine möglichst präzise Bestimmung der natürlichen Prognose, und die Frage lautet dann, «verändert meine Intervention den vermuteten natürlichen Verlauf?». Wir erkennen oder erinnern uns, dass auch beim Experiment der randomisierten kontrollierten Studie eine gleichartige Akzentverschiebung stattfindet: Eine sorgfältige Wahl prognostischer Faktoren als Ein- und Ausschlusskriterien ist als Patientenmerkmal mindestens so entscheidend wie die klassische Diagnose.

Fehlerhaftigkeit apparativer «Surrogat»-Medizin

Eine markante Akzentverschiebung findet aber auch in bezug auf die Wahl der Zielvariablen bzw. Endpunkte statt. Es wird zunehmend erkannt, dass die aus der «Euphoriezeit» stammenden Surrogate wie z. B. die Anzahl Extrasystolen im EKG

[12], der Cholesterinspiegel [13], die Tumorgrosse [14], die Knochendichte [15] etc. sehr «weiche», wenn nicht gar irreführende Zielvariablen darstellen können. Einen besonders interessanten Fall stellt vielleicht die Zielvariable Tumorgrosse dar, die zu mannigfaltigen Trugschlüssen verführt hat. Lassen Sie mich dazu nur sagen, dass wir bei differenzierter Betrachtung der Diagnose Mammakarzinom beispielsweise von einer so grossen prognostischen Streubreite ausgehen müssen, dass eine Erkrankung mit dieser Diagnose in vielen Fällen als klinisch gutartig bezeichnet werden kann [7, 14]. Dies führt dazu, dass wir in der Praxis neben einem Nutzen der Früherkennung in einigen Fällen auch einen Nutzen der Späterkennung in zahlreichen anderen Fällen beobachten [16]. Dies ist vielleicht ein besonders eindrückliches Beispiel, welches auf die Notwendigkeit neuer nosologischer Systeme oder einer neuen Taxonomie (wie Feinstein es nennt [3]) hindeutet.

Schliesslich spricht es sich inzwischen auch immer mehr herum, dass die in der Praxis sehr beliebten diagnostischen Surrogate – mit anderen Worten: Labor- und apparative Befunde – viel «weicher» sind als immer geglaubt wurde. Apparative Resultate vermitteln Wahrscheinlichkeitsaussagen, die nur zusammen mit klinischen Kriterien überhaupt vernünftig interpretierbar sind [17]. Zeichen eines Mammakarzinoms im Mammogramm etwa bedeuten ohne zusätzliche klinische Verdachtsmomente in über 90% ein falsch-positives Resultat; ein typischer Ischämie-Befund im Belastungs-EKG ist bei einem jungen Mann, der zum Check-up kommt, in gut 95% falsch-positiv. Diese Liste lässt sich fast beliebig verlängern.

Und hier kommen wir auf den springenden Punkt: Die akademische Medizin am Ende des 20. Jahrhunderts, so wie sie sich in Praxis und Lehre noch mehrheitlich darstellt, hat sich durch die «verführerischen Surrogate» technischer Möglichkeiten von ihrem Gegenstand *und* einer praktischen Rationalität so weit entfernt, dass sie uns Ärzte in der Praxis zunehmend im Stich lässt [18]. Dem sorgfältigen Beobachter fällt dabei auf, dass davon schulmedizinische Praktiker genauso betroffen sind wie Praktiker, die auf unkonventionelle Therapien ausgewichen sind. Einen Allgemeinpraktiker haben etwa die schwierigen Fragen der Multimorbidität, der Langzeitfolgen und der konkurrierenden Mortalität immer beschäftigt. Neue technische Errungenschaften im Computerzeitalter erlauben nun neue Forschungsfragen, weil Datenauswertungen und eine Logistik der Datenerhebung in ganz anderen Dimensionen möglich geworden sind.

Ich habe in meiner Nicaragua-Zeit beispielsweise die Masernimpfungen mit Begeisterung mitgemacht; die drastische Reduktion der in Dritte-Welt-Ländern häufigen Todesfälle durch Masern in diesem Land war ein schönes Ergebnis. Erst später ist mir in den Sinn gekommen, dass ich eigentlich wissen müsste, ob diese erfolgreiche Impfkampagne auch die Kindersterblichkeit insgesamt reduzieren konnte. Denn ein Kind, welches an Masern stirbt, ist von vornherein durch Unter-

ernährung kränklich; wenn es eine Maserninfektion nicht überleben kann, durch die Impfung nun aber von dieser verschont wird, wird es dann die nächste Grippe überleben? Wir freuten uns an der schönen Masernstatistik, es ist aber denkbar, dass wir durch die Impfung nur die Todesursache, nicht aber die Sterberate insgesamt beeinflussen konnten. Mit den heutigen Möglichkeiten der Datenverarbeitung werden solche Fragen aktuell.

Diese dramatische Wandlung der Forschungsfragen lässt sich auch am Beispiel Cholesterin zeigen. Die Informationsverarbeitung erlaubte zunächst unter grossem Aufwand Vergleichsstudien zwischen Ländern, dann die Beobachtung von grossen Kohorten, dann die Eruiierung der als cholesterinsenkend vermuteten Diäten, schliesslich einige grosse randomisierte Interventionsstudien. Dabei mussten die Nutzhypothesen ständig den neuen Möglichkeiten angepasst werden; schliesslich wurden grosse kontrollierte Interventionsstudien für einen Wirkungsnachweis als notwendig anerkannt und konnten dank moderner Datenverarbeitung bewältigt werden. Hier geht es aber noch weiter: Die Frage beschränkt sich heute nicht mehr nur auf die Wirkung der Cholesterinsenkung auf die koronare Herzkrankheit allein, sondern auf die Morbiditäts- und Mortalitätsentwicklung insgesamt, denn die Datenverarbeitungstechnik erlaubt schliesslich die zusammenfassende Auswertung der vielen Einzelstudien (Meta-Analysen). Erst jetzt konnte die den zu Ende denkenden Praktiker immer interessierende Frage nach der Gesamtmortalität beantwortet werden. Wir wissen jetzt, dass die medikamentöse Senkung des Cholesterins nur bei sehr hohem Gesamtrisiko einen günstigen Effekt hat, bei niedrigem Risiko die Sterblichkeit jedoch insgesamt erhöht [13].

Irrationalität in der Komplementärmedizin

Nun, wahrscheinlich haben Sie gemerkt, dass es nicht meine Absicht ist, hier eine Philippika gegen die naturwissenschaftliche Medizin zu führen. Wenn ich bisher hauptsächlich über mangelnde Rationalität, über Trugschlüsse und Herausforderungen in der offiziellen Medizin gesprochen habe, so ist dies nur möglich, weil innerhalb der Schulmedizin ein Prozess in Gang gekommen ist und weil gerade innerhalb der Schulmedizin neue beobachtungsorientierte Forschungsinstrumente von innovativen Köpfen geschaffen worden sind.

Vertreter von Komplementärmedizin scheinen demgegenüber oft mehr in alten Denkschemen verhaftet zu sein. So beschäftigte sich etwa der Akupunktur-Kongress in Rom im Jahre 1992 [19] fast nur mit der Frage, wie der Wirksamkeitsnachweis von Akupunktur mittels elektromagnetischen Surrogatmessungen oder etwa Akupunkturpunkte mit Hautdruckmessungen dokumentiert werden könnten. Der Schrei nach «harten» Messdaten bewegte diesen Kongress, offenbar in völliger Unkenntnis der diesbezüglichen Ergebnisse moderner klinisch-epidemiologischer Analysen und der wachsenden

Ernüchterung über das Dogma der harten Fakten unter modernen Wissenschaftlern. Daneben wurden Disputationen zelebriert über die richtige Auslegung der Bibeln der chinesischen Medizin.

Der hier kurz skizzierte methodische Wandel muss unbedingt zur Kenntnis genommen werden, wenn wir uns über die Frage nach der notwendigen Forschungsmethodik in der Komplementärmedizin unterhalten. Dieser Methodenwandel ist einerseits eine Chance für die Komplementärmedizin, seine Kenntnisnahme zwingt die Komplementärmedizin aber auch, ihre Wirksamkeit mit modernen Methoden zu dokumentieren. Es versteht sich dabei, dass weder Schulmedizin noch Alternativmedizin als allgemeingültige Glaubenssysteme zur Diskussion stehen, sondern es stehen bestimmte Therapien für bestimmte Indikationen zur Diskussion.

Zeitgemässe Wege der Forschungspraxis

Kommen wir zum Schluss aber noch zur Frage, welche methodischen Probleme sich etwa stellen bei der Planung einer Studie zur Wirksamkeit chinesischer Medizin – sagen wir bei Kopfschmerzen [20].

1. Problem: Definition der Studienpopulation

Eine Studienpopulation sollte einem klinischen Bild mit einer gewissen Homogenität entsprechen, welches in der Praxis wiedererkannt werden kann und für welches die untersuchte Behandlung wirklich sinnvoll erscheint. Für diese Gruppe sind dann die Ergebnisse aussagekräftig. Ich bin bereits ausführlich darauf eingegangen, dass eine klassische Diagnose – denken wir an Migräne, Spannungskopfschmerz, Cluster-Kopfschmerzen und posttraumatische Kopfschmerzen – gar nicht immer ein einheitliches Krankheitsbild umfasst. Wir müssen ebenso Ko-Morbidität, Dauer des Beschwerdebildes, Häufigkeit der Schmerzen etc. in Betracht ziehen. Betrachten wir Kopfschmerzen «chinesisch», dann gilt es folgende Unterscheidungen zusätzlich in Betracht zu ziehen: Ein nach oben schlagendes Leber-Yang, Hitze in Leber oder Magen, eine Wind-Bild, Feuchtigkeits- oder Schleimbelastung, Qi-Stagnation etc. (die Begriffe «qi» und Blut/«hsüe» lassen sich nach meiner Ansicht im übrigen am besten mit dem Begriffspaar «Kraft und Saft» übersetzen). Chinesische und westliche Kopfwehbilder überschneiden sich. Die Lösung des Problems liegt also darin, die Gruppe anhand von Einschluss- und Ausschlusskriterien praktisch sinnvoll zu definieren, d.h. unter Einbeziehung wichtiger klinischer und anamnestischer Merkmale aus beiden «Beobachtungssystemen». Eine gewisse Heterogenität der Untersuchungsgruppe kann und muss nicht gänzlich vermieden werden, oft wird vergessen, dass die unbekannt prognostischen Kriterien in der Regel die entscheidendsten sind. Randomisierung und statistische Testung sind dazu da, Vergleichbarkeitsmängel zu kompensieren und bürgen für die interne Validität einer Studie.

2. Problem: Definition der aktiven Behandlung und der Kontrolle

Bei einer Kontrolle denken wir oft zuerst an eine Placebobehandlung. Die praktische, wichtigere Frage ist zunächst aber die nach dem Stellenwert einer chinesischen Therapie im Vergleich zu einer konventionellen Therapie. Behandlungen in der Praxis sind konventionell oder chinesisch; Placebobehandlungen erfolgen kaum, und niemand behandelt mit Placebo-Akupunktur. So sind wir etwa auch innerhalb der konventionellen Medizin an Vergleichen zwischen internistischen oder chirurgischen Behandlungen interessiert; niemand käme dann auf die eher abwegige Idee, unbedingt Placebochirurgie als Kontrolle in Betracht zu ziehen, obwohl auch operative Eingriffe zweifellos Placebowirkungen haben. Eine Studie hat nur eine Validität, wenn die verglichenen Behandlungen bzw. Behandlungssysteme in der Praxis so existieren. Wenn wir also aus unbedachter Sorge um «Wissenschaftlichkeit» übermässig standardisieren, sinkt der wissenschaftliche Wert, denn dieser kommt durch die Praxisnähe der untersuchten Behandlung zustande. Eine mögliche Placebowirkung ist nicht von primärem Interesse; wenn ein Kopfschmerz und mögliche assoziierte Beschwerden sich bessern, ist ein Patient zufrieden und kümmert sich wenig um die eher akademischen Placebosorgen. Wichtig ist dem Patienten vielmehr eine anhaltende Besserung, d.h. entscheidend ist eine adäquat lange Beobachtungszeit (man kann es auch so sagen: Wenn mir Kollegen vorhalten, meine Akupunktur sei doch nichts anderes als Placebo, pflege ich zu sagen: Ja, ich kann das nicht sicher ausschliessen, doch mach mir doch diese Placebowirkung nach, dem Patienten geht es besser!). Wirklich wichtig und entscheidend für eine gute Studie sind hingegen eine möglichst weitgehende Verblindung der Beteiligten, eine Vermeidung von Erwartungshaltungen und Kontaminationseffekten (die Patienten in den verschiedenen Studienarmen sollten sich nicht sehen und gegenseitig beeinflussen können) sowie ein vollständiges Follow-up.

3. Problem: Die Definition der Zeitvariablen bzw. des Behandlungserfolges

Gott sei Dank, möchte man sagen, sind Kopfschmerzen nicht apparativ oder im Labor messbar, und die Gefahr der Wahl fraglich valider Surrogatmessungen ist hier gering. Als sinnvolle Zielkriterien kennen wir die subjektiven Schmerzen etwa anhand von Visual-Analogue-Scales sowie deren Frequenz, die Menge eingenommener Schmerzmittel und das Ausmass der Beeinträchtigung im Alltag. Wenn wir «chinesisch» behandeln, sind wir aber auch an der Entwicklung der Ko-Morbidität interessiert, denn es ist eine häufige Beobachtung, dass auch Randsymptome sich bessern, wie etwa Müdigkeit, Schlafstörungen, eine Hypertonie, chronische Nasen- und Nebenhöhlenentzündungen etc. Interessant und wichtig ist auch die Frage der Langzeitmorbidität und zwar der Multimorbidität. Es besteht Anlass zur Vermutung, dass eine erfolgreiche chinesische Behandlung auch eine präventive Wirkung

auf multimorbide Folgestörungen hat. Für die Praxis ist dies zweifellos eine wichtige und faszinierende Hypothese. Nur eine gute Dokumentation dieser präventiven Wirkung könnte der Akupunkturmedizin gegenüber der doch bequemeren konventionellen Medikamenteneinnahme einen wichtigen Stellenwert in der Behandlung von Kopfschmerzen verschaffen.

Schlussfolgerung

Wir brauchen keine alternative Forschungsmethodik für die Komplementärmedizin. Ich hoffe, dass ich zeigen konnte, dass Allgemeinmedizin und Komplementärmedizin vielmehr vor denselben methodischen Problemen stehen. Es gilt, die konventionellen methodischen Entwicklungen der letzten Jahre, die recht tiefgreifend sind, zur Kenntnis zu nehmen und damit kreativ umzugehen.

Literatur

- 1 Smith P: Paracelsus as emblem. *Bull Hist Med* 1994; 68:314–322.
- 2 Kochen MM: Forschung in der Allgemeinmedizin; in Kochen MM (Hrsg): *Allgemeinmedizin*. Stuttgart, Hippokrates, 1992, pp 519–529.
- 3 Feinstein AR: The hard data creed in current clinical practice: Its spurious validity and the challenge to define meaningful clinical variables; in Schmidt JG, Steele RE (Hrsg): *Kritik der medizinischen Vernunft: Schritte zu einer zeitgemässen Medizin – Ein Lesebuch*. Mainz, Kirchheim, 1994; pp 210–218. (Deutsch: Das irreführende Credo harter Daten in der heutigen klinischen Praxis – das Problem sinnvoller klinischer Erfolgskriterien. *Forsch Komplementärmed* 1994;1: 260–265.)
- 4 Hislop TG et al: The prognostic significance of psychosocial factors in women with breast cancer. *J Chron Dis* 1987;40:729–735.
- 5 Ramirez AJ et al: Stress and relapse of breast cancer. *Br Med J* 1989;298:291–293.
- 6 Spiegel D et al: Effect of psychosocial treatment on survival of patients with metastatic breast cancer. *Lancet* 1989;ii:1209–1210.
- 7 Schmidt JG: The epidemiology of mass breast cancer screening: A plea for a valid measure of benefit. *J Clin Epidemiol* 1990;43:215–225,235–239.
- 8 Syme SL: Control of destiny and health: Towards a more effective preventive medicine; in Schmidt JG, Steele RE (Hrsg): *Kritik der medizinischen Vernunft: Schritte zu einer zeitgemässen Medizin – Ein Lesebuch*. Mainz, Kirchheim, 1994, pp 100–112.
- 9 Feinstein AR: Clinical judgment revisited: The distraction of quantitative models. *Ann Intern Med* 1994;120:799–805.
- 10 Gästrin G: Preliminary results of primary screening for breast cancer with the Mama Program. *Soz Präventivmed* 1993;38:280–287.
- 11 Gästrin G et al: Incidence and mortality from breast cancer in the Mama Program for Breast Screening in Finland, 1973–1986. *Cancer* 1994;73:2168–2174.
- 12 Schmidt JG: Antiarrhythmika bei Extrasystolie? *pharma-kritik* 1991;13:41–44.
- 13 Schmidt JG: Der Einfluss der Cholesterinsenkung auf die Mortalität. *Schweiz Med Wochenschr* 1993;123:2150–2155.
- 14 Schmidt JG: Natural history of breast cancer. *Lancet* 1992;339:810.
- 15 Bucher HC, Schmidt JG: Osteoporose-Prophylaxe. *pharma-kritik* 1994;16:13–16.
- 16 Schmidt JG: Früherkennung und Umgang mit Risikofaktoren; in Kochen MM (Hrsg): *Allgemeinmedizin*. Stuttgart, Hippokrates, 1992; pp 133–149.
- 17 Bucher H, Gutzwiller F: *Checkliste Gesundheitsberatung und Prävention*, Stuttgart, Thieme, 1993.
- 18 Schmidt JG: Erfolgsnachweis in der Allgemeinmedizin. *Z Allg Med* 1992;68:720–724.
- 19 WFAS International Symposium on the Trend of Research in Acupuncture (World Federation of Acupuncture-Moxibustion Societies). Abstracts & papers. Rome, 22–24 October 1992.
- 20 Schmidt JG: Measuring the efficacy of acupuncture medicine; in Schmidt JG, Steele RE (Hrsg): *Kritik der medizinischen Vernunft: Schritte zu einer zeitgemässen Medizin – Ein Lesebuch*. Mainz, Kirchheim, 1994, pp 203–207.