

V. Die Vision einer pragmatischen klinischen Forschung: Methodologie und Studiendesigns / The Vision of Pragmatic Clinical Research: Methodology and Study Designs

Kommentar des Herausgebers

Die Leistungsfähigkeit von Medizin ist ein Produkt aus der Wirksamkeit ihrer Mittel einerseits und der breiten Anwendbarkeit und Generalisierbarkeit ihrer wirksamen Mittel andererseits. Fehlt die Wirksamkeit der Mittel, wird auch bei grosser Generalisierbarkeit dieser Mittel eine geringe Leistungsfähigkeit resultieren. Sind andererseits die Mittel sehr wirksam, aber nur bei einem kleinen Spektrum von Gesundheitsstörungen eine adäquate Antwort, resultiert ebenfalls eine geringe Leistungsfähigkeit. Die derzeitige «schulmedizinische» Forschung leidet unter dieser Beschränktheit, nur ein kleiner Bruchteil heutiger klinischer Studien sind methodisch valide und gleichzeitig relevant für die Praxis [1]. Vermutlich ist und bleibt der Leistungsauftrag an die Medizin, *immer zu trösten, oft zu lindern und gelegentlich zu heilen* [2], was sich immerhin im heutigen Wechsel vom Seziertisch-Endpunkt zur Lebensqualität als Leistungskriterium zu widerspiegeln scheint.

Damit ist auch die Leistungsfähigkeit klinischer Forschung, praxisrelevante Antworten auf reelle Probleme zu geben, herausgefordert. Um die Vision einer pragmatischen klinischen Forschung nachzuvollziehen, braucht man indessen kaum hellseherische Fähigkeiten, denn die erforderliche Methodologie und entsprechende Studiendesigns sind durch die klinische Epidemiologie und die Klinimetrie bereits solide erarbeitet [2, 3], was auch im folgenden gezeigt werden kann. Klinische Studien müssen «intern valide» sein, und eine randomisierte Kontrollgruppe muss deshalb wenn immer möglich hergestellt werden (natürlich nach adäquater Präzision der Forschungsfrage mit Hilfe nichtkontrollierter Studien), damit nicht immer wieder die notorische Selektion erfolgreicher Fälle einen Behandlungserfolg vortäuscht, der in Wirklichkeit nicht vorhanden ist. Die Leistung klinischer Studien hängt aber ebenso davon ab, dass Therapievergleiche praxisgerecht und wenn nötig individualisiert erfolgen und dass sie die «komplexen» Patienten einschliessen, die sich in der Praxis finden. Die grosse Aufgabe klinischer Forschung ist heute, Orientierung zu schaffen im Gestrüpp der vielen Heilsversprechen – alternativmedizinischer und schulmedizinischer. Pragmatische klinische Studien sind das Mittel dazu (auch wenn sie weniger präzise Antworten auf die oft

Editorial Commentary

The performance capacity of medicine is a product, on the one hand, of the effectiveness of its means and, on the other hand, of the broad applicability and generalisability of its effective means. If the means are not effective, then low-quality performance results even if the generalisability of these means is high. If, conversely, the means are very effective, but only provide an adequate solution in a very small range of health disorders, then similarly low-quality performance results. Currently, research in 'classical medicine' is suffering from this latter limitation. Only a small fraction of today's clinical studies are methodologically valid while being clinically relevant at the same time [1]. Presumably, the job medicine is supposed to fulfil is and remains *to comfort always, to relieve often and to cure occasionally* [2], which indeed appears to be reflected by the current transition from the autopsy table endpoint to overall quality of life as the relevant outcome measure.

This challenges the ability of clinical research to provide clinically relevant solutions to real problems. One hardly has to be a clairvoyant to comprehend the vision of a pragmatic clinical research since the required methodology and the corresponding study designs are already soundly based with the existence of clinical epidemiology and clinimetrics [2, 3]. Clinical studies must be 'internally valid' and a randomised control group must therefore be created whenever possible (naturally after adequate precision of the research question has been established from observational studies), so that the notorious selection of successful cases does not keep mimicking a therapy success that is not there. The performance of clinical trials, however, is equally dependent on the inclusion of therapeutic operations that are practice-oriented and, when necessary, done on an individualised basis and that they include the 'complex' patients who are found in routine practice. The major task of clinical research today is to provide orientation in the dense thicket of the many promises of a cure offered by both alternative as well as by conventional medicine. Pragmatic clinical studies are the means of doing so (even if they produce less precise answers to the often rather academic questions of specific agent theories).

eher akademischen Fragen spezifischer Wirkungstheorien geben können).

Bereits im ersten Beitrag leistet *Roger A. Edwards* eine ausgezeichnete, zeitgemässe Übersicht über Ziele und Methoden klinischer Studien und erklärt die von Feinstein stammenden Konzepte einer «fastidious efficacy» und einer «pragmatic efficacy». *Edwards* macht klar, dass die Wahl der Studienmethodik immer von intelligenten Kompromissen abhängt, die konkurrierende Anliegen möglichst optimal zur Geltung zu bringen haben. Er macht auch klar, dass es letztlich darum geht, möglichst effizient Fehlermöglichkeiten («chance», «bias» und «confounding») zu kontrollieren, und dies ist bei genügend Sorgfalt bei gewissen Fragestellungen auch in nichtrandomisierten Studien möglich. Wichtiger als die falsche Frage eines alleingültigen Designs ist für den Fortschritt der klinischen Forschung vielmehr die sorgfältige Formulierung der Forschungsfragen. *Edwards'* methodisch differenzierte Darlegung verschiedener Studiendesigns ist trotz Relativierung der randomisiert kontrollierten Studie als Standard des Wirksamkeitsnachweises indessen keine Annullierung der Hierarchie der Evidenz (vgl. Kapitel III). Dem genauen Stellenwert der Randomisation geht der Beitrag der Biometriker *Koch und Abel* noch eingehender nach. Dabei machen sie klar, dass Randomisation alleine noch kein Gütesiegel für eine Studie ist. Denn die Randomisation kann zwar Selbstselektion der Patienten und selektive ärztliche Gruppenzuordnung kontrollieren, garantiert aber nicht, ob die Therapie in beiden Armen gleich motiviert und konsistent durchgeführt worden ist und die Ergebnisse auf gleiche Weise beurteilt worden sind. Randomisation ist auch überhaupt keine Garantie dafür, dass die geprüfte Therapie Sinn macht und die klinisch relevanten Endpunkte gewählt worden sind. Wenn die Medizin – und hier ist insbesondere auch die Komplementärmedizin angesprochen – aufhört, aus so vielen unsorgfältigen Beobachtungsstudien «Kausalität» und «Erfolg» abzuleiten, dann wird die Diskussion über den Stellenwert methodisch sorgfältiger Beobachtungsstudien ohne randomisierte Kontrollen mehr Platz bekommen können [4]. *Jos Kleijnen* geht auf die genaue Rolle der Placebo-Kontrolle ein, wobei er diese Frage auf das klinisch enge Spektrum der Überprüfung neuer Medikamente beschränkt. *Kleijnen* begründet die Notwendigkeit einer Placebo-Kontrolle aus der Beobachtung, dass der «informed consent» und die daraus geweckten Erwartungen die Wirkung stark beeinflussen können, vor allem die Nebenwirkungen. Er macht dadurch klar, wie kritisch die Ergebnisse über neue, «bessere» Medikamente betrachtet werden müssen und dass eine Placebo-Kontrolle (in einem drei- oder vierarmigen Design) über den Vergleich mit der etablierten Therapie hinaus entscheidende Informationen liefert, vorausgesetzt eine solche Kontrolle ist überhaupt möglich und es besteht eine entsprechende Frage. *Edzard Ernst* beleuchtet den Stellenwert des «n = 1 trial» und von Fallserien. Auch diese Designs sind für gewisse Fragen unersetzlich, haben aber gleichzeitig ihre Grenzen. Die externe Validität des «n = 1 trial» ist auf den untersuchten Patienten beschränkt; Fallserien sind immer einer Reihe von Bias unterworfen, so dass sie eine Therapiewirksamkeit in der Regel nie zweifelsfrei

Already in the first contribution of this chapter, *Roger A. Edwards* provides us with an excellent, timely review of the goals and methods of clinical studies and explains the concepts of a 'fastidious efficacy' and a 'pragmatic efficacy' originating from Feinstein. *Edwards* made it clear that the choice of study methodology always depends on intelligent compromises that must take competing interests into account as optimally as possible. He also clearly states that the ultimate aim is to control biases (chance and confounding) as efficiently as possible, and, if done carefully enough this is possible for certain issues even in non-randomised studies. Careful formulation of research questions is much more important for achieving progress in clinical research than the secondary question of what constitutes a universally valid study design. Despite relativising the randomised clinical trial as an exclusive standard for demonstrating efficacy, *Edwards's* methodologically differentiated presentation of various study designs, in the end, does not negate the hierarchy of evidence (cf. chapter III). The exact place of randomisation is examined further in the report by the biometricians *Koch and Abel* in which they make clear that randomisation is no guarantee at all for the validity of a study. Although randomisation can control self-selection of patients and biased group allocation by doctors, it is no guarantee for a consistent treatment motivation in both arms and for a consistent outcome assessment. Moreover, randomisation cannot guarantee that the therapy makes sense and that clinically relevant endpoints have been selected. Once medicine – and here I especially refer to complementary medicine – stops deriving 'causality' and 'success' from so many carelessly designed observational studies, then there will be more room for the discussion about the place of methodologically well-designed observational studies without randomised control [3]. *Jos Kleijnen* discusses the role of placebo control while limiting this question to the clinical narrow spectrum of testing new drugs. *Kleijnen* justifies the need for placebo control based on the observation that 'informed consent' and the expectations arising therefrom can greatly influence the effects, particularly the side-effects. In doing so, he illustrates how important it is to view the results on new, 'better' drugs critically and that a placebo control (in a three- or four-arm design) can provide important information that goes beyond comparison with the established reference therapy; this is providing that such a control is possible at all and that the corresponding information is needed. *Edzard Ernst* highlights the value of the 'n = 1 trial' and case series. Here, too, these designs are indispensable for examining certain questions, but also have their limitations at the same time. The external validity of the 'n = 1 trial' is limited to the investigated patient; case series are always subject to a number of biases so that they can rarely document the efficacy of a therapy with certainty. *Peter Mattmann-Allamand* explains the way homoeopathy works and shows that because of the nature of its mechanisms it is scientifically wrong to test homoeopathic monopreparations. On the other hand, the important question as to what effect a homoeopathic treatment setting might have in certain indications compared to a conventional therapeutic procedure can, under certain circumstances, be tested very effectively in pragmatic studies. His own study of homoeopathy during pregnancy is

dokumentieren können. *Peter Mattmann-Allamand* erläutert die Arbeitsweise der Homöopathie und zeigt, dass es aufgrund der Natur dieser Arbeitsweise wissenschaftlich falsch ist, homöopathische Einzelmittel zu prüfen. Die wichtige Frage, was eine homöopathische Arbeitsweise bei gewissen Indikationen im Vergleich zu einem konventionellen Therapievorgehen bringen kann, lässt sich hingegen in pragmatischen Studien unter Umständen recht gut überprüfen. Seine eigene Studie über Homöopathie in der Schwangerschaft stellt ein sehr gutes Beispiel für einen pragmatischen Vergleich dar, auch wenn die Studie an ihrer zu kleinen Grösse «gescheitert» ist. Auch wenn *Mattmann* vom «indirekten» Wirksamkeitsnachweis spricht, ist die «pragmatische Wirksamkeit» individualisierter homöopathischer Therapie genau das, was wir direkt wissen wollen. Wenn sich die «direkte» Wirksamkeit für den Therapeuten selbst etwa an der sogenannten Erstverschlimmerung ablesen lässt, so bleibt ja die entscheidende Frage, ob die ausgelösten «Wirkungen» dem Patienten am Schluss einen Nutzen bringen. Die zu Recht aufgeworfene Frage, ob sich Patienten überhaupt von einer homöopathischen Therapie in eine Kontrollgruppe mit konventioneller Therapie «weg-randomisieren» lassen, zielt indessen weniger auf die Randomisierung als solche, sondern auf die wichtige Frage, was wann randomisiert werden soll. Wenn das therapeutische Arbeitsbündnis und/oder das weltanschauliche Einverständnis mit der Homöopathie möglicherweise eine entscheidende Rolle spielt, dann gilt es, die Randomisierung so zu legen, dass diese Interaktionen nicht auseinanderdividiert werden. Vielleicht ist die praktisch-klinisch entscheidende Frage oft die nach der Wirksamkeit von Homöopathie bei gesellschaftlichen Untergruppen. Eine pluralistische Medizin, die erkennt, dass viele Wege nach Rom führen können, wird eine insgesamt wirksamere Strategie in der Hand haben. *Karl-Ludwig Resch* stellt im letzten Beitrag dieses Heftes die vielfältigen Möglichkeiten dar, wie Randomisationen erfolgen können. Randomisierte Kontrollen können (bei genügend Erfindergeist) auch für komplexe Therapien bzw. Therapiestrategien hergestellt werden, auch wenn Behandlungen individualisiert erfolgen. Das nicht nur von Vertretern komplementärmedizinischer Verfahren und von Psychotherapeuten, sondern auch von konventionellen Allgemeinpraktikern oft vorgebrachte Argument, jeder Patient sei einzigartig und deshalb liessen sich gewisse Therapien gar nicht in Gruppenvergleichen untersuchen, ist letztlich ein Denkfehler. «Erfahrungen» aus der Praxis sind bei genauer Betrachtung immer «statistische Auswertungen» der eigenen Fälle, die allgemeingültigere Gesetzmässigkeiten erkennen lassen; *Resch* weist zu Recht darauf hin, dass aus Behandlungserfahrungen gar nichts gelernt werden könnte, wenn jeder Fall einzigartig wäre. Pragmatische klinische Studien sind bisher rar. *Resch* nennt die wichtige Existenzbegründung für das «3. Wissenschaftliche Einsiedler Symposiums» und dieses ganze Heft, wenn er darauf hinweist, dass die Weiterentwicklung pragmatischer klinischer Studien die wichtigste Herausforderung an die klinische Forschung der Zukunft ist.

Pragmatische klinische Studien erlauben die bitter nötige Wirksamkeitsforschung für komplementärmedizinische Verfahren und

a good example of a pragmatic comparison, even though it ‘failed’ because of its small size. Although *Mattmann* speaks of an ‘indirect’ demonstration of efficacy, the ‘pragmatic efficacy’ of individualised homoeopathic therapies is exactly what we directly want to find out. Although the ‘direct’ efficacy itself is determined by the homoeopathic doctor based on the so-called initial worsening, the crucial question still remains as to whether the ‘effect’ that has been triggered benefits the patient. The legitimate question whether patients would let themselves be randomised away from a homoeopathic therapy into a control group receiving conventional therapy, if scrutinized carefully, is less directed at randomisation per se, but rather at the important question as to what should be randomised when. If the therapeutic working alliance and/or the ideological consent to homoeopathy does play an important role, then randomisation should be designed in such a way that it does not separate these interactions. Perhaps, the key question for clinical practice is often that of the effectiveness of homoeopathy in social subgroups. Pluralistic medicine that acknowledges that many roads lead to Rome will be in possession of an overall more effective strategy. In the last contribution here, *Karl-Ludwig Resch* presents the myriad possibilities of how randomisation can be conducted. Randomised controls (using a bit of inventiveness) can also be created for complex therapies or therapeutic strategies, even if the treatments are given on an individualised basis. The argument often presented not only by advocates of complementary medical procedures and psychotherapists, but also by conventional general practitioners that each patient is unique and therefore subgroup comparisons cannot be used at all to test certain therapies ultimately constitutes an error in reasoning. Upon close examination, ‘practical experience’ is always a ‘statistical analysis’ of one’s own cases which produces a generally valid regularity; *Resch* rightly points out that nothing could be learnt at all from therapeutic experience if every case were truly unique. Pragmatic clinical studies have been rare so far. *Resch* states the main reason for the existence of the ‘3rd Scientific Symposium of Einsiedel’ and for this entire issue when he points out that the further development of pragmatic clinical studies is the major challenge to clinical research of the future.

Pragmatic clinical studies allow the badly needed research into the efficacy of complementary medical procedures and give answers about the effect that a therapeutic procedure or a therapeutic school of thought has in a specific indication. The question as to which component actually constitutes that effect may remain open in such studies. Nonetheless, pragmatic clinical studies are also necessary in conventional medicine, where we not only need to find out about the efficacy of a drug, but also about the benefits of its use within the whole therapeutic setting. Here, too, comparisons of therapeutic strategies often provide much more important information than the effect of a single isolated therapeutic agent. Today, general practice in particular can rarely fall back on studies that have posed their questions and provided answers that are valid and useful in practice.

geben Antwort auf die Frage, welche Wirkung eine therapeutische Vorgehensweise oder eine therapeutische Schule bei gewissen Indikationen hat; die Frage, was genau die Wirkung ausmacht, darf in einer solchen Studie offen bleiben. Pragmatische klinische Studien sind aber auch in der konventionellen Medizin erforderlich, wo wir nicht nur über die Wirksamkeit eines Medikamentes, sondern über den Nutzen seines Einsatzes im ganzen therapeutischen

Kontext Bescheid wissen müssen. Auch hier sind Vergleiche therapeutischer Strategien oft eine weit wichtigere Information als die Wirkung eines isolierten therapeutischen Agens. Insbesondere die Allgemeinmedizin kann heute selten auf Studien zurückgreifen, die ihre Fragen gestellt haben und praktisch gültige Antworten geben.

Literatur/References

- 1 Field MJ, Lohr NN (eds): Institute of Medicine Guidelines for Clinical Practice: From Development to Use. Washington DC, National Academy Press, 1992.
- 2 Feinstein AR: Why do we need clinical epidemiology? A practice-oriented clinical science. In: Schmidt JG, Steele RE (eds): Kritik der medizinischen Vernunft: Schritte zu einer zeitgemässen Praxis – Ein Lesebuch. Mainz, Verlag Kirchheim, 1994, pp 233–243.
- 3 Feinstein AR: Clinimetrics. New Haven, Yale University Press, 1987.
- 4 Abel U, Koch A (eds): Nonrandomized Comparative Clinical Studies. Düsseldorf, Symposium Publishing, 1998.