Forschende Komplementärmedizin
Research in Complementary Medicine

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

Forsch Komplementärmed 1998;5(suppl 1):116–120

Roger A. Edwards

Harvard Medical School, Santa Fe, NM, USA

Research and the Goal of Improving Patient Care

Key Words

Outcome research · Randomized controlled trials · Research evaluation

Summary

The goal of this article is to raise issues associated with the underlying assumptions of our research approaches and the nature of health/disease in conventional medical care and complementary and alternative medical care. Different types of efficacy are discussed in terms of randomized controlled trials and outcomes research. The demands of different audiences for research results suggest that researchers need to carefully specify the research question in the context of the state-of-the-knowledge. Appropriate matching of the study design, the nature of the disease, the nature of the treatment and what is known is essential to designing and implementing research that actually could improve patient care.

Schlüsselwörter

Outcome-Studien · Randomisiert kontrollierte Studien · Forschungsevaluation

Zusammenfassung

Ziel der Forschung ist die Verbesserung der Leiden von Patienten

Ziel dieser Arbeit ist es, die unseren Forschungsansätzen zugrundeliegenden Annahmen über die Natur von Gesundheit/Krankheit in der konventionellen und der komplementären und alternativmedizinischen Versorgung zu untersuchen. Verschiedene Arten von Wirksamkeit sowie deren Evaluation in randomisiert kontrollierten Studien und in Verlaufsstudien werden diskutiert. Die verschiedenen Interessen an Forschungsergebnissen erfordern eine sorgfältige Formulierung der Forschungsfrage. Die Übereinstimmung von Studiendesign, Art der Erkrankung, Art der Behandlung mit dem Kenntnisstand ist eine wesentliche Voraussetzung für eine Forschung, die tatsächlich die Patientenversorgung verbessert.

Does our research really achieve the goal of improving patient care? I would like to begin with the consideration of the question if we really have theory behind our research. The origin of the word 'theory' comes from the Greek word *theoria*, which means 'to behold' [1]. Theories provide frameworks for conceptualizing and analyzing phenomena. Ideally, the process of gathering empirical data to test new theories and the formulation of new theories to explain empirical data works to advance our understanding of ourselves and our realities. In the scientific sense of the word, the essence of progress is the willingness to move bidirectionally between theory and empiricism toward the goal of understanding the nature of the universe. Each of us has the opportunity to contribute to the evolution of science of health and health services by evaluating the theories implicit in the assumptions we retain about health and health services research and the assumptions we behold about the mystery of healing – and ensuring that in the end these assumptions meet the overarching goal we share: the improvement of patient care.

I would like to introduce some concepts that relate to health and healing and that are not new, and I want to ask a question: Do we want to maintain the dominant paradigm or do we want to adopt a new one? If so, what is the new paradigm going to consist of? There is a lot of discussion about wholeness, balance, energy, physical well-being, mental well-being, spiritual well-being – but what really is health and is it peculiar to an individual? There is research in the U.S. that asks people to pick the three things that are of greatest importance to them in their lives and then to measure the health intervention's impact on these things, so that we can begin to understand how unique the perception of health is for an individual and measure what is important to that individual. It has been phrased by one individual as 'the precision of modern science with the wisdom of ancient healing' [2]. Does our research and do our

Fax (07 61) 4 52 07 14
www.karger.com
Accessible online at:
http://BioMedNet.com/karger

Roger A. Edwards, ScD
Instructor, Harvard Medical School
Research Consultant, Partners Telemedicine
Center/Massachusetts General Hospital
2918 Plaza Blanca, Santa Fe, NM 87505 (USA)
Tel. 505 473-96 03, Fax -34 44

research methodologies really reflect these concepts? Sometimes they do and sometimes they don't. So that goes back to the question: Why are we doing the research? Are we really developing 'systematic investigations designed to develop and contribute to generalized knowledge' [3]? What did it really produce for the investment? I refer to complementary and alternative medicine (CAM). Are our CAM studies making such a contribution? Is our conventional medical research making such a contribution? Then we have to ask, who are the consumers of our research? As researchers we often talk among ourselves. But different audiences would accept different approaches. Physicians or other providers might accept one set of approaches. Payers might accept another set of approaches and patients a different one. I think, whatever method we adopt, we have to ask if we are documenting it in such a way that it is reproducible. Reproducibility is the hallmark of good science. And yet few of us can pick up something in the literature and reproduce the study. Now, granted, that is partly a constraint of the journals and their page limits and so forth, but there is a policy that is evolving related to submitting of appendices corresponding to the articles that individuals can request to help address this issue of reproducibility in science.

One example of a multidimensional healing process is the one proposed by Dacher [4]. Dacher describes the reorientation of primary care to include an expanded healing model by going from the assumption of the biomedical model of objectivism, determinism, and positivism to dynamism, holism, and purposefulness [4]. This expanded paradigm emphasizes the quality and character of the relationships that contribute to the healing process. The consideration of the healing relationship from a number of perspectives is being undertaken by several foundations in the U.S., including the Robert Wood Johnson Foundation and the Fetzer Institute. They now have or will have requests for proposals (RFPs) that relate to this issue of the therapeutic relationship. So, clearly, the 'relationship' is becoming more and more important to consider in the healing process. If we do integrate this concept with the existing paradigm, then what do we have? Does our research reflect that? I think it speaks to an issue raised many years ago by Sir William Osler: 'It is better to know the patient that has the disease than the disease that has the patient' [4]. Maybe in the last forty years we have focused on the latter, and maybe we want to consider the former. Certainly, there has been a backlash in the U.S. related to this issue of patients not being considered as whole human beings in the health care process.

So what do we have available to us to consider in terms of 'does something work or does something not work'? My colleagues, Ted Kaptchuk and David Eisenberg, and I [5] tried to pull together the literature on what is effectiveness and what is efficacy. Essentially, the framework had three different aspects of efficacy or effectiveness. In the first case, it is fastidious efficacy, and this term came out of Alvan Feinstein's work [6]. Fastidous efficacy assumes that the verum specific effects can be added to the nonspecific effects to get the total outcome effect. It conceptualizes discrete components that are additive, and it is the underlying premise behind the randomized controlled trial (RCT). There is the double-blind randomized control trial which has been used by regulators on several continents for decades to try to insure that safe and efficacious medications actually come on the market. And to that end, it has been quite a successful tool because when you are dealing with powerful pharmaceuticals, the issues of whether it can be safe and be efficacious at the same time were very important. We have all benefited as citizens by the use of a technique of this sort in this application. The question is how and when should it be applied, and in what situations should it not be applied?

In addition to that there has been a growing movement related to pragmatic efficacy [7], otherwise thought of as 'effectiveness'. Now this theoretical framework states that you cannot separate the specific effects from the nonspecific ones. They are linked in such a way that they cannot be separated; therefore, you can only look at the total outcome effect. Sometimes they may be linked in a synergistic way, sometimes in an antagonistic way; but either way they cannot be separated, and the only way you can understand outcomes is to look at something in actual clinical practice under actual clinical situations in contrast to the fastidious efficacy framework which seeks to look at something from an idealized situation where you control as many variables as you can ethically can control for and try to determine the cause and effect relationship. Essentially, the fastidious framework tries to maximize internal validity issues, and the pragmatic framework tries to optimize external validity. Each has limitations, and neither goes as far as it could.

There is another framework that comes out of the medical anthropology literature and was developed by Tambiah [8]; it states that one should look at performative efficacy. What he means by performative efficacy is more cultural and anthropological. It accepts that symbols, beliefs, suggestions, expectations, and persuasions are central to illness and health, and it is often implicit in the qualitative research approaches which really have not been discussed to a great extent here. In this case, what matters is the meaning to the patient.

Now, these three paradigms, these three theories, underlying 'what works', coexist, are relevant in different stages of knowledge, and are not mutually exclusive. I think part of the debates that we have had are artificial and unnecessary because they try to pit these against each other when really they all go together depending upon the state of the knowledge.

The issue of the randomized control trial, particularly the double-blind one, has widely been discussed. The question raised about 'double blinding' or 'randomization' being the gold standard is interesting. I was always taught that it did have to be double blind to be the gold standard, but that may vary by educational organizations.

I would like to briefly go into the question why you would look at observational research methods and what can be found in the literature about these approaches. One of the first ones is ethics. We have heard a lot of discussion about ethics, but there is also one other aspect touched on in almost every epidemiological textbook that I have seen. There is an appropriate time to do an RCT, and

that is when the existing therapy and the new therapy, based upon the knowledge gathered to date, are equivalent in terms of not knowing which is better. Would you put yourself in such a trial? Would you put a loved one in such a trial? And if the answer is 'no' – that you prefer one therapy to the other – then maybe that is not the right time to do an RCT. I have heard physicians stand up in conferences and say they would never put a patient into this trial because they always have some knowledge that would indicate one therapy over another. To me that presents an ethical dilemma, and I think we need to ask ourselves very seriously of the trials that we are doing: 'Are they really done at the appropriate stage?'

A common example is the study of aspirin to prevent heart attacks that had physicians as subjects. Already some physicians had started to take aspirin related to prevention of heart attacks. The researchers had that as a confounder when they first started the study because the state of the knowledge was just beginning to become evident that you wanted to take aspirin. They got the study concluded in time before that confounder was a major factor, but it spoke to the issue of when you do an RCT from an appropriate ethical perspective.

The other aspect is the feasibility of the intervention and the nature of complex health problems. I will not belabor these points, but want to point out that some situations are very hard to study from a double-blind point of view and also from a randomized point of view.

The other issue is cost. It is extremely expensive to do a double-blind, randomized control trial. It does not mean you should not do it, but maybe we as a society have to ask: 'How do we want to invest our resources?' What is the appropriate cost/information tradeoff? What about the stage of evolution of the knowledge? It is not just the knowledge of the intervention, it is the knowledge of the disease we are studying, and often that fact is very hard to incorporate into our decision making, the review of the literature, the clinical experience, and so forth. This phase of reviewing all these issues, of synthesizing and thinking them through is often done very quickly and without sufficient thought; I think it is really important to spend as much time as possible on this phase.

In addition, you have quality no matter what kind of design you have; you have prospective studies that tend to be better than retrospective ones, and you might have random sampling of providers from lists of licensure, for example, before you begin an observational study to try to improve the quality of an observational study and begin to blend the best of each design.

The other aspect is that observational studies can often help with hypothesis generating for when you get ready to do an RCT down the road and also after something has been approved.

Another big use of observational studies, in the United States at least, has been improvement in quality of care – particularly as it relates to continuous quality improvement. In this sense the goal has been to reduce unnecessary variation and preserve systematic variation, the difference being that unnecessary variation potentially harms the patients. I think the movement and the work in this area has done a very good job of bringing up the lower threshold, so that people are no longer getting the worst care. What has been harder to achieve is the goal of improving care when the choice is from among fairly effective therapies beyond a certain threshold. That is a harder task, and a lot of the work with practice guidelines has been important in this area. But with practice guidelines the issue remains which patients are appropriate for the guidelines and if the provider has sufficient flexibility to opt out of the guideline when the patient situation demands that.

I mentioned generalizability already with regard to a broader patient population. In many cases you are seeing payers who are demanding effectiveness studies. They are saying we want to see how it will work in our population; we want to see the outcomes and the cost implications of those outcomes. They want the RCTs, but they also want the well-done, prospective observational studies, and so forth. Finally, you are also beginning to see a recognition of the patient/provider dynamics.

Obviously, some of the limitations of observational studies relate to the reliability of the data and the quality of the input. In the cases of retrospective studies the data were often collected for other purposes; financial purposes and key medical information, particularly related to the severity of the illness, are often lacking. Therefore, we have to question the reliability of the data – whether it was collected prospectively or retrospectively – and also the validity in terms of the comparability of patients across situations and also coding. The coding of the information may alter or bias what is happening. I think this is one of the weaknesses that is a limitation, unless you can really begin to address that issue from a number of perspectives.

The other issue is patient confidentiality. Who should have access to information? What kind of information? In the European continent usually people will not be denied health care because of their health information. That is not true in the United States. So it becomes a very important question as to who has access to what information because it could lead to denial of patient access to care at a later point.

I would like to return to the provider/patient dynamics and research design from different perspectives. The fastidious approach assumes that we can alter the relationship. It is appropriate to mitigate the healing effects of a trusting provider/patient relationship in an RCT. Is the greater good worth the tradeoff, worth the loss in a particular study situation? We, as a society, say that yes, some individuals can be denied the benefit of an individual therapeutic relationship because the rest of society will benefit from that knowledge. Certainly, many individuals who consent to be in medical research think they are helping knowledge and patients subsequent to themselves when they do agree to participate.

The pragmatic approach says that it is not appropriate to separate the specific and nonspecific effects and, therefore, you should not disrupt the provider/patient relationship. It should stay intact. Therefore, an observational approach of what is going on in practice – an effectiveness study – is useful. I think that the evidence is so sparse on this issue of the tradeoffs of disrupting or not disrupting the provider/patient relationship that we really don't know.

White et al. [9] advocate a systems theory approach related to the placebo that had 22 interrelated elements. There is a lot of empirical work that could be done in this area concerning the interactions among these 22 and other elements that relate to this provider/patient relationship. Assuming that we accept the benefits of randomization, we must be aware that we are only randomizing patients, and patients are part of the healing process. What if we need to randomize providers? Maybe we should randomize the match or the pairing between the patient and the provider because an authoritarian provider will do very poorly with a patient who wants to be empowered; whereas a doctor or physician who empowers would do very well with a patient who wants to be empowered, but not very well with a patient who wants to be told what is best for him or her. So, if we are randomizing patients for the familiar arguments, maybe we also need to randomize the other important elements in the whole health-care process including the providers and the provider/patient pairings. But would providers consent to this? Would patients consent to this? And if they would not consent – why? Do we need to reconsider the whole approach?

One alternative that tries to bring together some of the strength of the different designs and minimize the weaknesses, although it has its flaws too, is a preference trial. A preference trial was introduced into the literature, as far as I could find, by John Wennberg back in around 1990 [10]. Essentially, a preference trial says that you would want to randomize patients to one of two conditions, a 'choice of a therapy» or 'assignment of a therapy'. If a patient is randomized to 'choice of therapy', then he or she could chose therapy x or therapy y based upon a script, upon an interactive video disk, upon any amount of information that is provided related to the benefits, the costs, and the risks associated with those two therapies. On the other hand, the other patient that was randomized through 'assignment of therapy' would follow the traditional design of being assigned therapy x or y. In summary, you can compare therapy x when it was chosen or when it was assigned and you might be able to find out if the outcome was different as a result of the choice. In that case you might see a choice effect or a preference effect. There is no reason to believe that the choice effect would be constant. Some therapies might have a high choice effect, and some therapies might have a low one. Some would argue that in the case of our very most powerful pharmaceuticals, the choice element is very low; the choice effect would be very low because it is such a powerful intervention. Setting a broken bone might be another example. On the other hand, in the case of more subtle therapies with mechanisms that are less well understood, the choice effect might be higher, but we don't know. It is an empirical question that is yet to be studied. But at least with this kind of design, you might see the magnitude of the choice effect in two different kinds of therapies. You also might be able to see if this mattered in whether therapy x or therapy y is better. If the difference could be accounted for by choice, than this might say something else as well.

I would like to close with consideration of this issue of evidence in clinical studies. First of all, there are not homogeneous categories. They could have hybrid outcomes – RCT studies and you can have double-blind, single-blind studies, etc. This issue has been discussed to quite an extent. The catch is that all of these studies, no matter what the design, are concerned about chance, bias, and confounding. So when thinking about the studies and reviewing the literature, it is important that we, as researchers, address the role of chance, the role of bias, and the role of confounding no matter which method we choose. I think we lose sight of some of that, and it is important to think about it. One should remember that none of these methods is sufficient alone. We need all depending upon the stage of knowledge regarding a particular disease and the stage of that knowledge regarding a particular intervention. Each can be enriched by the other. For example, the Framingham study is a longitudinal cohort study that has been going on for approximately forty years in the United States of a community and its health, particularly related to cardiovascular disease. That observational study has generated a tremendous amount of information related to risk factors and the modeling of those risk factors. Many individuals, including the regulators, asked that the results of an RCT be used as inputs into the Framingham data – into the Framingham models. The RCTs are more useful because of the observational study. The observational study is more meaningful because you can put in state-of-the-art RCT data. So each is better off, and certainly we as citizens who are trying to optimize our healing and health care are benefited by the interaction and synergy of the two types of methods.

I think it is time we put aside this artificial dichotomy between RCTs vs. observational studies. I think it is moving in this direction. In my opinion the really basic issue is the formulation of the research question. We often overlook this because the same design can be really great for one research question and really poor for another and we kind of scoot past this because it is very tedious and time consuming to specify it, to do a design and then go back and see how we really address the research question that we set out to do. And does that research question reflect for which audiences we are doing the research? Because different audiences have different questions. Furthermore, the issue of the difficulty of interpreting the placebo response is fascinating to researchers and perplexing to providers. In my view we need to think about this issue in terms of our designs and who we are doing them for.

I think we can put aside our ideological debates and let RCT imperialism and outcomes research imperialism die and be done with, and even consider new approaches. One approach that is getting a lot of attention is the Bayesian approach, which tries to combine subjective perception of what is going on and quantifies, given the limitations in that quantification process, with objective evidence. I think the reanalysis of the GUSTO trial from 1995 [11] was an example of what would happen if we adopted a Bayesian approach to this issue.

I would like to raise one last issue. Physics went through a shift of paradigms and experimental approaches, and it took decades and decades for physicists to deal with the move from classical mechanics to quantum mechanics. Health care and health services research have yet to go through such a transformation. Maybe it should and maybe it shouldn't, but I think we have to ask ourselves when it does, if our frameworks do better fit – behold – the phenomena that we are studying.

References

1 Zayon A: Catching the Light. New York, Oxford University Press, 1993, p 22.
2 Gordon I: Center for Mind-Body Medicine, Washington DC.
3 Pak CYC, Adams PM: Techniques of Patient Oriented Research. New York, Raven Press, 1994, p 36.
4 Dacher ES: Reinventing primary care. Alt Ther Health Med 1995;1:29–34.
5 Kaptchuk TJ, Edwards RA, Eisenberg DM: Complementary medicine: efficacy beyond the placebo effect; in Ernst E (ed): Complementary Medicine: An Objective Appraisal. Oxford, Butterworth-Heineman, 1996.
6 Feinstein AR: Clinical Epidemiology. Philadelphia, W. B. Saunders, 1985.
7 Schwartz D, Lellouch L: Explanatory and pragmatic attitudes in therapeutic trials. J Chronic Dis 1967;20:637–648.
8 Tambiah SJ: Magic, Science, Religion, and the Scope of Rationality. Cambridge, Cambridge University Press, 1990.
9 White L, Tursky B, Schwartz G (eds): Placebo: Theory, Research, and Mechanisms. New York, The Guilford Press, 1985.
10 Wennberg JE: What is outcomes research? in Gelijns AC (ed): Modern methods of Clinical Investigation. Washington DC, National Academy Press, 1990, pp 33–46.
11 Brophy JM, Joseph L: Placing trials in context using Bayesian analysis: GUSTO revisted by Reverend Bayes. JAMA 1995;273:871–875.