

E. Ernst

Department of Complementary Medicine,
Postgraduate Medical School, University of
Exeter, UK

Establishing Efficacy in Chronic Stable Conditions: Are 'N = 1 Study' Designs or Case Series Useful?

Key Words

Complementary medicine · Research methodology · N = 1 study ·
Case series · Validity

Summary

The aim of this overview is to discuss the usefulness of two research tools often advocated in complementary medicine: the n = 1 study and case series. These methodologies are defined and their advantages and disadvantages are outlined. It is concluded that both designs have advantages and disadvantages. For testing the efficacy/effectiveness of complementary therapies neither design will lead to conclusive, generalizable results. Yet both methodologies can be valuable adjuncts to other types of investigation.

Schlüsselwörter

Komplementärmedizin · Untersuchungsmethodik · N = 1-Studie ·
Fallstudien · Validität

Zusammenfassung

Wirksamkeitsnachweis bei chronisch stabilen Erkrankungen: Bringen «N = 1»-Studien oder Fallserien etwas?

Ziel dieser Übersicht ist es, die Verwendung zweier in der Komplementärmedizin oft verwendeter Untersuchungsmethoden zu diskutieren: die «N = 1»-Studie und Fallstudien. Diese Methoden werden definiert und ihre Vor- und Nachteile skizziert. Es zeigt sich, dass beide Methoden Vor- und Nachteile haben. Soll die Wirkung bzw. Wirksamkeit komplementärmedizinischer Therapien untersucht werden, ergibt keine der Methoden schlüssige, verallgemeinerbare Ergebnisse. Allerdings können beide Methodologien als Ergänzung für andere Arten von Untersuchungen nützlich sein.

Introduction

Throughout the history of medicine, it has been assumed that any positive change in symptomatology following a therapeutic encounter is most likely due to the therapy administered. In short, the therapy has been given the benefit of the doubt. But this is not necessarily correct. There can be many reasons for clinical improvements; and if we ascribe them to the treatments we will create a 'false-positive' result. The recognition of this phenomenon in the 1940s led to the development of a range of controlled clinical studies culminating in the randomized clinical trial (RCT). Essentially controlled trials aim at eliminating bias from unrelated factors to allow specific therapeutic effects of treatment to be indubitably shown. Today the RCT (not, as often claimed in complementary medicine, the double-blind trial or the placebo-controlled trial) has therefore become the 'gold standard' for testing efficacy of medical treatments. Of course, most experts recognise the drawbacks of this methodology. They principally are due to the fact that

trial conditions are 'artificial situations', and that there may be ethical problems inherent in appointing patients to groups. The RCT may be the best available but is by no means the optimal method, and it is legitimate, even necessary, for medical scientists to improve their methods. The following article will briefly summarize the advantages and disadvantages of 'n = 1 studies' and (even more briefly) of 'case series' as a means of establishing efficacy of medical interventions.

'N = 1 Studies'

Characteristics

'N = 1' studies are prospective experimental studies with a sample of one. On this single subject or unit a treatment is sequentially given or withheld, and the change in an outcome variable (e.g. pain score) is measured during this repeated process [1]. This methodol-

ogy is not new; it was developed for physiological research in the 1970s [2] and was adopted by medicine about 10 years later [3]. Many variations of the basic theme exist [4]. Inter alia, 'n = 1 studies' can be

- randomized or not
- single/double-blind or not
- placebo-controlled or not
- comprised of one, two or many cross-overs.

They can also be conducted on a group of individuals (e.g. one family) as long as they represent one unit for the purposes of data analysis. In some ways, such studies are similar to conventional cross-over trials: The number of cross-overs may vary and therefore a whole range of treatments can be tested and re-tested in a single study. The fact that 'n = 1 studies' are, by definition, experimental clearly differentiates them from descriptive single case studies in which there is no repeated intervention or indeed no pre-defined study protocol.

Advantages

The results of conventional clinical trials provide us with nothing more than probabilities. Frustratingly, we therefore never know whether the next patient will benefit from the experimental treatment. 'N = 1 studies' were developed not least to overcome this drawback. If well-designed with a sufficient number of cross-overs to guarantee intra-individual reproducibility, they are capable of achieving this aim [3]. The 'n = 1 design' is therefore attractive for treatments in which a range of ingredients are tailored closely to individual needs, e.g. homeopathic prescription. However, the design is not the only one attempt to overcome the problem of highly individualized treatments [e.g. 5]. Even the standard parallel group design can achieve this aim [e.g. 6].

It has been claimed [1] that 'n = 1 trials' avoid the ethical dilemma of treatment denial to patients recruited for a clinical study. This is only partly true. Certainly, 'n = 1 studies' require no control group (which might receive placebo), but there are still 'ethical hot spots' during treatment when the single subject receives no active therapy. Certainly, it is the case that in the course of a trial the patient invariably receives *active* treatment but this is also true for conventional cross-over designs.

A major and obvious advantage of the design is the fact that only one subject is required. 'N = 1 studies' can therefore be performed in conditions which are too rare for conventional trials (and for which therefore no trial data exist). Unquestionably, this advantage has to be set against the disadvantage of lack of generalisability (see below).

Disadvantages

'n = 1 studies' are problematic for conditions which are not stable, e.g. chronic conditions of constant severity or with repeated episodes of similar severity. Rather than relying on inter-group com-

parisons (as most conventional trials) they compare points in time (e.g. condition without treatment vs. condition with treatment) in single patients or units. Such comparisons are based on the assumption that there is no baseline-shift or instability in symptoms. If, however, the condition investigated changes over time (e.g. through the natural history of the disease), the results of such comparisons are hard to interpret. As most conditions in medicine do change with time, and 'stability' is in fact a continuum, the applicability of the 'n = 1' design is limited (at least to a some degree) in many clinical situations. Instability can be made worse by the fact that a given treatment changes over time, e.g. becomes less effective, a phenomenon known as tachyphylaxia.

An even more important disadvantage of this design is its lack of external validity. In other words, even if the study arrives at definitive conclusions, its results are applicable only to the individual tested. One way of overcoming this shortcoming is to conduct a number of 'n = 1 studies' large enough to be representative for an entire patient population [7]. Such approaches can yield results with reasonable generalisability; they are, however, similar to conventional cross-over trials. One may therefore well ask, why not opt for the latter approach in the first place? The question becomes even more pertinent when the conclusions from pooled 'n = 1 studies' state that 'larger studies ... are needed to find definite answers' [8].

'N = 1 trials' can often not be evaluated by simple test statistics which often assume independence of data points. More over, they are prone to carry over effects (persistence of therapeutic effect into next treatment phase) and period effects (difference of response between treatment phases, e.g. due to a saturation of some biological response). They therefore require complex statistical approaches like time series analyses [9]. The obvious disadvantage is that the expertise for such analyses is not always available and that their results are not always comprehensible to ordinary healthcare professionals.

Like all trials, 'n = 1 studies' are prone to type I and II errors, i.e. false-positive or false-negative results. One way of minimizing them, is to increase the number of cross-overs, i.e. the number of times treatment is stopped or started. It has been suggested that 20 or more treatment periods are needed unless treatment is very effective [10].

'Case Series'

'Case series' are accumulated case reports evaluated either retrospectively or (more rigorous) prospectively. They can vary in quality (e.g. have better defined inclusion/exclusion criteria, more sensitive endpoints etc). 'Case series' (similar terms: cohort studies, uncontrolled observational studies) may seem attractive to many investigators as they do not require informed consent, pose no problem in terms of treatment denial, and fit comfortably into clinical settings. But their most important methodological drawback is the lack of a control group. Thus they have no place in the evaluation of clinical efficacy: Their results simply do not tell us whether an observed

change was indubitably due to the therapy or to any of the following factors, each of which can influence the clinical results [11]:

- placebo effect
- natural history of the disease
- regression to the mean
- patient's desire to please the doctor
- doctor's desire for a positive result
- concomitant therapy
- other non specific effects.

This, however, is not to say that 'case series' are of no value. They are certainly useful in a number of ways. Often they are even essential for formulating a hypothesis [12]. In turn, this hypothesis, however, requires testing by other methods, e. g. RCTs.

Conclusions

'N = 1 studies' can bridge the gap between research and clinical practice [13]. They can provide definitive efficacy data for individual patients, particularly for treatments already investigated conventionally [14] though generalisability from one case is poor. Thus they are not an alternative to other methodologies but can be a useful adjunct. 'Case series' are valuable tools, for instance, for hypothesis generation but inadequate for testing efficacy.

References

- 1 Sim J: The ethics of single-system (n = 1) research. *Physiother Theory Pract* 1994;10:211–222.
- 2 Kratchwill TR (ed): *Single Subject Research: Strategies for Evaluating Change*. Orlando, FL, Academic Press, 1978. p 316.
- 3 Guyatt G, Sackett D, Wayne Taylor D, Chong J, Roberts R, Pugsley S: Determining optimal therapy – randomized trials in individual patients. *N Engl J Med* 1986;314:889–892.
- 4 Hinkel J, Hornung J: 'Single-Case Studies' – Eine Übersicht. *Forsch Komplementärmed* 1997;4:6–15.
- 5 Ernst E, Resch KL: The 'optional cross-over design' for randomized controlled trials. *J Fundam Clin Pharm* 1995;9:508–511.
- 6 Reilly D, Taylor MA, Beattie NGM, et al: Is evidence for homoeopathy reproducible? *Lancet* 1994;344:1601–1606.
- 7 March L, Irwig L, Schwarz J, Simpson J, Chock C, Brooks P: N of 1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. *BMJ* 1994;309:1041–1046.
- 8 Mahon J, Laupacis A, Donner A, Wood T: Randomised study of n of 1 trials versus standard practice. *BMJ* 1996;312:1069–1074.
- 9 Edington ES: Statistics and single case analysis. *Prog Behav Modif* 1984;16:83–119.
- 10 Lewis JA: Controlled trials in single subjects. II. Limitations of use. *BMJ* 1991;303:175–176.
- 11 Ernst E, Resch KL: The clinical trial – gold standard or naive reductionism. *Eur J Phys Med Rehabil* 1996;6:26–27.
- 12 Vickers A, Cassileth B, Ernst E, Fisher P, Goldman P, Jonas W, Kang S-K, Lewith G, Schulz K, Silagy C: How should we research unconventional therapies? A panel report from the conference on Complementary and Alternative Medicine Research Methodology, National Institutes of Health. *Int J Technol Assess Health Care* 1997;12:111–121.
- 13 Johannessen T: Controlled trials in single subjects. I Value in clinical medicine. *BMJ* 1991;313:173–174.
- 14 Campbell MJ: N of 1 trials may be useful for informed decision making. *BMJ* 1994;309:1045.