Editorial

Wien Klin Wochenschr (2005) 117/7–8: 241–242
DOI 10.1007/s00508-005-0343-x

WIENER KLINISCHE WOCHENSCHRIFT
The Middle European Journal of Medicine
Printed in Austria

On the inconclusiveness of "evidence"

The article on anthroposophic treatments in this issue of *Wiener klinische Wochenschrift* is a welcome occasion to remember a few essential points about the conclusiveness or otherwise of clinical trial evidence [1]. Most of us will know the "hierarchy of evidence" within evidence-based medicine [2]. Figure 1 depicts this hierarchy in the shape of a pyramid. The basis of the pyramid is formed by all types of uncontrolled data. These can be case reports, case series, observational studies or even history of usage (anthroposophic treatments, for instance, have been popular for decades). I used the term "basis" intentionally – such data are indeed the foundation for much in clinical medicine. In particular, they are invaluable for formulating hypotheses. However, if we want to test hypotheses, they are usually next to worthless. For testing hypotheses, we need controlled clinical trials and, if we want to minimize bias in such trials, we must randomise. Even randomised studies often generate conflicting results, and therefore systematic reviews of such trials are often the least biased way of determining whether a given treatment is effective or not.

The article by Hamre at al. in the current issue describes a controlled multi-centre trial of patients receiving either anthroposophical or conventional treatment for infections [1]. The results show that the former group experienced more and faster benefit than the latter group. Anthroposophical treatment was also associated with more patient satisfaction and with less adverse effects. All in all a very positive result – should we therefore use anthroposophic treatment for such indications? I'm not sure.

One very fundamental weakness of this study prevents it from generating conclusive results that can be translated into clinical practice: patients were self-selected. That means they or their parents (many patients were children) had chosen to see either an anthroposophical or a conventional doctor. Therefore the two groups differed in several known variables and perhaps in many other unknown variables as well. Some of these differences can go a long way to explain the findings of the study. Sadly such biases render trials of this nature more valuable to marketing than to science. Here are only four possible, perhaps even likely, explanations why the anthroposophical patients had better results than the conventional ones and four conclusions which are consistent with the study results:

– The anthroposophical group contained more children, and children might generally recover quicker from infections than adults. Conclusion: infections have a different natural history in children compared to adults.
– Conventional doctors may have prescribed antibiotics unnecessarily, and we all know that this can have detrimental effects, particularly on the frequency of adverse events. Conclusion: unnecessary intake of antibiotics is detrimental to health.
– The anthroposophical patients suffered less frequently from difficult to treat conditions such as sinusitis and therefore recovered quicker. Conclusion: the nature of disease affects its course and outcomes.
– Anthroposophic treatment "entails more active engagement" of patients (as the authors state). This could create a more powerful placebo effect in the anthroposophic compared to the conventional group. Conclusion: patients' involvement in therapy can affect the outcome.

Any of these circumstances suffices to generate a false positive result. In other words, the outcome could be unrelated to the specific treatment. The authors' attempts to control for confounding mathematically may have been not fully successful and, of course, can only extend to known variables. What about factors that we don't know about?

The inconclusiveness of this trial is annoying because, no doubt, the study will now be interpreted as evidence supporting the efficacy of anthroposophical medicine (the authors state that "Anthroposophical treatment is … at least as effective as conventional treat-

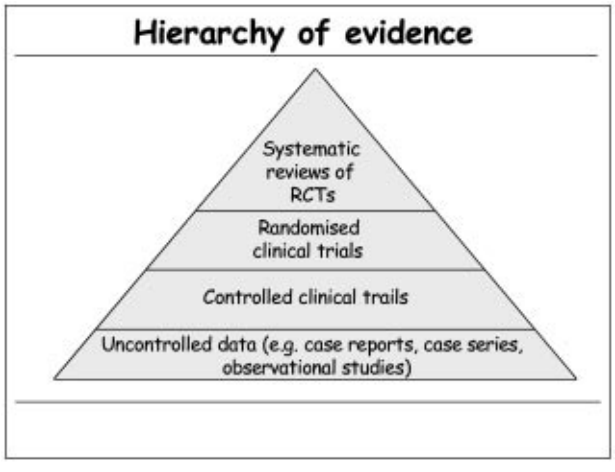

Fig. 1

ment"). It is also avoidable – simply by using randomisation instead of self-selection. The fact that "randomisation has traditionally been rejected in anthroposophical medicine" is hardly a convincing argument.

Some people would argue in defence of the trial that pragmatic studies cannot be as rigorous as other trials; the flaw of selection bias must therefore be accepted. This, however, is both incorrect and nonsensical. Pragmatic trials can certainly be designed such that selection bias is minimised [3]. Trials which use pragmatism as an excuse for inconclusiveness are not pragmatic but worthless in answering the research question posed. In the present case, the objective was "to compare anthroposophic treatment to conventional treatment". I fear that this objective cannot be achieved with the study design chosen by the authors.

Clinical trials are about finding the truth. If they are severely biased they can seriously mislead us. And misleading results are, of course, a risk factor for good health. Clinical trials also need to be reproducible and valid. Anyone attempting to reproduce the present study would run into considerable difficulties. For instance, the diagnoses of "respiratory and ear infections" are vaguely based on clinical judgment. And outcome measures such as "complete recovery" or "major improvement" are subjective and not validated. Anthroposophical physicians participating in this study obviously knew what was at stake; it is conceivable that their judgements were clouded by their hope for a "positive" result. In my view, fatally flawed studies can be more than just a waste of resources and opportunities; they can be worse than no evidence at all.

The moral of this story is simple. If anyone claims that their therapy is effective, they must deliver the evidence. This evidence has to comply with any currently accepted standards of scientific rigour. Double standards cannot be tolerated because they are counter-productive and, in the extreme, endanger the health of our patients.

Edzard Ernst

References

1. Hamre HJ, Fischer M, Heger M, Riley D, Haidvogl M, Baars E, Bristol E, Evans M, Schwarz R, Kiene H (2005) Anthroposophic vs. conventional therapy of acute respiratory and ear infections: a prospective outcomes study. Wien Klin Wochenschr 117: 256–268
2. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS (1996) Evidence based medicine: what it is and what it isn't. BMJ 312: 71–72
3. Vickers AJ, Rees RW, Zollman CE, McCarney R, Smith CM, Ellis N, et al (2004) Acupuncture for chronic headache in primary care: large, pragmatic, randomised trial. BMJ 328: 744

Key words: Evidence, clinical trials, pragmatic trials.

Correspondence: Edzard Ernst, MD PhD FRCP FRCPEd, Complementary Medicine, Peninsula Medical School, Universities of Exeter & Plymouth, 25 Victoria Park Road, Exeter EX2 4NT UK, E-mail: Edzard.Ernst@pms.ac.uk