Special Section

THE QUALITY OF THE MEDICAL EVIDENCE: IS IT GOOD ENOUGH?

Section Editors

Kay Dickersin
Andrew Herxheimer
INTRODUCTION: THE QUALITY OF THE MEDICAL EVIDENCE: IS IT GOOD ENOUGH?

Kay Dickersin  
*University of Maryland at Baltimore*

Andrew Herxheimer  
*The U.K. Cochrane Centre*

No one who opens this special issue will expect the answer “yes” to this question: it is very far from good enough. This is despite the fact that thousands, if not millions, of reports are published every year in over 16,000 biomedical periodicals. It seems that although much research has been done, it is not the “right” research. It is also possible, perhaps probable, that the right research has been done, but it cannot be accessed by those who need it or has not been noticed.

The central problem is that much research is not very good. Studies that test health care interventions very often have biased or unreliable results. For example, it is not unusual to see a comparison of two groups of patients who received two different forms of care: one form of care routinely given by Doctor A and the other by Doctor B. However, it would be unwise to compare the effectiveness of the two forms of care received using these two populations, since the doctors and their patients are likely to differ: the doctors may attract patients from different socioeconomic groups, or one of the doctors may tend to have the most unusual and most severe cases referred to her. Thus, a selection bias may be operating that would affect the findings of any study conducted where assignment to intervention was not randomized.

Another shortcoming of many studies testing the efficacy of an intervention is the failure to include enough participants. Small samples result in less precise estimates of effect and may fail to identify truly effective interventions. Both selection bias and small sample size have long been recognized as common study defects, yet on average, most clinical trial reports continue to have these and other problems and to be of fairly low quality.

This special issue of the Journal addresses a variety of concerns of practical importance to determination of the efficacy of health care interventions:

- What is the best way to define and assess the quality of research?
- Does the source of funding influence the quality and findings of research studies?
- Are we measuring effectiveness using appropriate outcome measures?
- In what circumstances should we worry about bias in the analysis and reporting of research?
Dickersin and Herxheimer

• How can the needs of patients and ethical considerations be satisfied and how may these considerations affect study quality?
• How can we ensure that the available evidence is collected systematically, completely, and kept up to date?
• Is the research community focusing too much on the quality of the evidence, while ignoring consumers', health care providers', and policy makers' need for more evidence?

Paul Stolley sets the scene by reminding us how shaky the foundations of clinical practice have been, even during our lifetime. His entertaining account makes a very serious point.

Even though randomized controlled trials are the most reliable source of evidence on the value of an intervention, not all are equally good. David Moher and his colleagues examine the thorny problems of assessing their quality and review the various approaches that have been tried. Whether the holy grail of a single numerical score to indicate the quality of a study will turn out to be a mirage remains to be seen.

Studies undertaken for commercial reasons have long been particularly suspected of bias, but such studies have only recently been systematically examined. Lisa Bero and Drummond Rennie describe the extent to which these suspicions are justified. Their thorough exploration of the many aspects of this issue provides many interesting examples.

The understandable impatience of sponsors and investigators to find quick answers has led to the widespread use of surrogate outcome measures instead of the longer-term outcomes that matter to patients. Peter Gøtzsche and his colleagues warn us of the dangers of surrogates, giving examples where wrong conclusions long delayed the correct evaluation of important therapeutic interventions.

In the design of trials it is often difficult to strike the right balance between conflicting considerations, such as the desire to use random allocation to minimize bias and the desire to allow patients to choose the treatment that they prefer. Judith Lumley, working in perinatal medicine, and Hilda Bastian, a consumer advocate, consider the various conflicts that arise and cite numerous current examples.

A problem with published trial reports that has only recently been recognized stems from the fact that even a long and detailed paper can give only a summary of the data. Because space in paper journals is limited, much potentially important information about the trial inevitably remains unpublished and unknown. Lesley Stewart and Max Parmar discuss the many ways in which bias may enter into the conduct, analysis, and reporting of trials, and what reviewers can do to try to detect and correct such biases. Everyone who has to evaluate treatments will appreciate what they say.

Systematic reviews are now accepted as an essential part of the evaluation of health care. The international Cochrane Collaboration is creating, maintaining, and disseminating systematic reviews of health care for all areas of medicine. The type of work that is involved in producing high-quality systematic reviews for the Cochrane Collaboration is described by Murray Enkin and Jini Hetherington, who pioneered it in the area of pregnancy and childbirth.

Finally, Chris Hyde, from his standpoint and experience as a public health physician developing policy from research results, explains that many obstacles have to be overcome for research findings to be translated into practice. The quality of evidence is only one of these hurdles.
We hope that this special issue of the *International Journal of Technology Assessment in Health Care* will be useful as a guide to aid research, reviews, teaching, and policy development in health care. It should help health professionals and consumers to recognize good studies, and authors and editors to eliminate reports that are uninformative and cloud our views about what “works.”