What kind of randomized trials do we need?

Merrick Zwarenstein MD MSc, Shaun Treweek BSc PhD

Early release, published at www.cmaj.ca on Apr. 16, 2009. This article was published simultaneously in the May 2009 issue of the Journal of Clinical Epidemiology (www.jclinepi.com).

In 1967, Daniel Schwartz and Joseph Lellouch, 2 French statisticians, and their British colleague and translator Michael Healy wrote “[M]ost therapeutic trials are inadequately formulated, and this from the earliest stages of their conception.”

The seminal paper1 from which this dramatic assertion is drawn is reprinted in the May 2009 issue of the Journal of Clinical Epidemiology as part of a joint focus with CMAJ on making randomized controlled trials (RCTs) more useful.

Schwartz and Lellouch argued that there are 2 kinds of randomized trials embodying radically different attitudes to evaluation of treatment, which they named “pragmatic” and “explanatory.”2 They go on to say that these 2 attitudes require different approaches to the design of a randomized trial. The pragmatic attitude seeks to directly inform real-world decisions among alternative treatments. Schwartz and Lellouch show that this purpose is satisfied in trials that select typical participants, settings and comparator care to widen real-world applicability. In contrast, the explanatory attitude seeks to understand a biological process by testing the hypothesis that the specified biological response is explained by exposure to a particular treatment. Tight restrictions on participants, treatment, control and setting maximize the contrast with the control group and increase the ability to test this kind of hypothesis.

Their assertion of inadequate formulation relates to the mismatch between the use we make of most trials (which is to inform decisions on therapy) and the design of these trials (which generally takes the opposite form, best suited to testing causal hypotheses). This mismatch between the clinical context in which clinicians must make decisions and the clinical context of the randomized trials that they must use for evidence means that health professionals (and, writ large, health care funders) are left without clear evidence upon which to base most of the patient care decisions they make. Since information from an explanatory trial is unlikely to inform a pragmatic question, Schwartz and Lellouch proposed that investigators should explicitly specify the purpose of their trial and design it to match that purpose.

There are few trials whose purpose and design choices match. The only review of this subject2 identified fewer than 100 pragmatic designed randomized trials, of the quarter million or so RCTs listed by the US National Library of Medicine, which suggests that existing RCTs are mostly explanatory in design and thus not directly applicable to choosing between treatment options. This is ironic since the very first published randomized trial was pragmatic in purpose and in many of its design choices. It showed clear benefits for patients receiving streptomycin and usual care (bed rest) over the control group receiving only usual care.3 The decline of tuberculosis in high-income countries is thus due in part to the pragmatic trial.

Why so few pragmatic trials? Because of the size of the market, US pharmaceutical licensing regulations are the principal stimulus for the conduct of RCTs of treatments and the main influence on their design. The requirement that pharmaceutical manufacturers demonstrate efficacy of their products was first legislated in the 1962 Kefauver–Harris amendments to the US Federal Food, Drug, and Cosmetic Act, passed in the wake of the thalidomide tragedy, from which the United States had been largely spared by caution on the part of the US Food and Drug Administration (FDA).4 Perhaps because of their focus on safety, these licensing regulations5 devote most of their attention to preparatory studies in animals, safety issues in humans and documentation. The little guidance there is argues against trials with a pragmatic attitude: “One problem [with active-control trials] is that there are numerous ways of conducting a study that can obscure differences between treatments, such as poor diagnostic criteria, poor methods of measurement, poor compliance, medication errors, or poor training of observers. As a general statement, carelessness of all kinds will tend to obscure differences between treatments. Where the objective of a study is to show a difference, investigators have powerful stimuli toward assuring study excellence.”6 Much of what the FDA labels “careless” or “poor” is typical in usual care. The FDA thus equates pragmatic design choices aimed at increasing applicability with carelessness and poor study design, which results in trials that lack the attributes needed to directly support decisions about the real-world usefulness of a treatment.

FDA regulations and guidance are influential, but 2 other factors cannot be ignored: after spending many millions of dol-
Commentary

Table 1: Key differences between trials with explanatory and pragmatic attitudes*

<table>
<thead>
<tr>
<th>Feature</th>
<th>Explanatory attitude</th>
<th>Pragmatic attitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question</td>
<td>Efficacy: Can the intervention work?</td>
<td>Effectiveness: Does the intervention work when used in normal practice?</td>
</tr>
<tr>
<td>Setting</td>
<td>Tightly controlled, well resourced, “ideal” setting</td>
<td>Normal practice</td>
</tr>
<tr>
<td>Participants</td>
<td>Highly selected. Poorly adherent participants and those with conditions that might dilute the effect are often excluded</td>
<td>Little or no selection beyond the clinical indication of interest</td>
</tr>
<tr>
<td>Intervention</td>
<td>Strictly enforced; adherence is monitored closely</td>
<td>Applied flexibly as it would be in normal practice</td>
</tr>
<tr>
<td>Comparator</td>
<td>Strictly enforced; adherence is monitored closely</td>
<td>Often usual care, with usual variation. Applied flexibly as it would be in normal practice</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Often short-term surrogates, or process measures</td>
<td>Directly relevant to participants, funders, communities and health care practitioners</td>
</tr>
<tr>
<td>Relevance to practice</td>
<td>Indirect: little effort is made to match the design of the trial to the decision-making needs of those in the usual setting in which the intervention will be implemented</td>
<td>Direct: the trial is designed to meet the needs of those making decisions about treatment options in the setting in which the intervention will be implemented</td>
</tr>
</tbody>
</table>

*Adapted, with permission, from Zwarenstein M, Treweek S, Gagnier J, et al.; CONSORT and Pragmatic Trials in Healthcare (Practih) groups. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008;337:a2390. The table in BMJ was adapted from a table presented by Marion Campbell, University of Aberdeen, at the 2008 Society for Clinical Trials.

...also appears in this issue of CMAJ (www.cmaj.ca/cgi/rapidpdf/cmaj.090076), points to the close match between decision-makers’ preferences for directly applicable evidence and the pragmatic attitude favoured by Schwartz and Lellouch. Some of these authors have also recently published an extension to the CONSORT (Consolidated Standards of Reporting Trials) statement for pragmatic trials (Table 1),14 with recommendations on reporting of trials whose aim is to inform decisions. This CONSORT extension encourages authors to include in their trial publications information to help readers judge the applicability of the results to their own settings, the better to decide whether or not to implement the tested intervention (Table 1).14

The expansion of interest in the design of pragmatic randomized trials to support decision-making is surely now much needed. In Canada, for example, the regulatory authority (Health Canada) is currently considering a progressive licensing model for therapeutic agents.15 Foremost in its consideration is speeding the process of licensing while maintaining high standards for the evaluation of safety and efficacy. Health Canada might consider revisiting its regulations and encouraging the use of pragmatic designs to support its decision-making. It might be argued that this would lengthen an already prolonged licensing process, reducing the ability of inventors to profit from their discoveries. There is a counter argument: if the initial randomized trial of a treatment were pragmatic in purpose and design, funders of the treatment could immediately use that information to make decisions on usefulness in their setting and patient group, eliminating what are today entirely separate and sequential processes for regulatory approval and formulary inclusion. Combining the trials that collect information for regulatory and for formulary inclusion processes could simplify both and make more transparent the reasons for decisions on reimbursement. And since large public funders demanding information for decision-making are the world’s most important markets for new phar-
maceuticals, this could lengthen the period of profitable patent protection, rather than shorten it.

Schwartz and Lellouch ended their paper with a devastating conclusion: “Most trials done hitherto have adopted the explanatory approach without question; the pragmatic approach would often have been more justifiable.” Forty years and hundreds of thousands of randomized trials later, this remains true. It is time to shift our design choices so that they match our usual purpose in conducting a trial, most often to directly inform the decisions of real-world patients, clinicians and third-party funders.

Competing interests: None declared.

Contributors: Both of the authors were involved in the preparation of the manuscript and approved the final version submitted for publication.

REFERENCES

Correspondence to: Dr. Merrick Zwarenstein, Centre for Health Services Sciences, Sunnybrook Research Institute, Sunnybrook Health Sciences Centre, Rm. G1-06, 2075 Bayview Ave., Toronto ON M4N 3M5; merrick.zwarenstein@ices.on.ca