ELSEVIER

Journal of Clinical Epidemiology 63 (2010) 1283–1284

Journal of Clinical Epidemiology

N-of-1 clinical trials should be incorporated into clinical practice

Jesse A. Berlin*

Johnson & Johnson Pharmaceutical Research and Development, 1125 Trenton-Harbourton Road, PO Box 200, M/S TE3-15, Titusville, NJ 08560, USA

Accepted 30 May 2010

Abstract

N-of-1 clinical trials have the potential to contribute to individual patient management and to the accrual of important information about populations. Incorporating these studies into clinical practice will require creative thinking so as to maintain rigor without excessive disruption of routine care.

Keywords: Clinical trial design; *N*-of-1 clinical trials; Meta-analysis; Personalized medicine; Comparative effectiveness

Discussions of comparative effectiveness, patient-centered research, and personalized medicine have become common in the medical literature, newspapers, and legislative settings in the United States. In that context, the article by Zucker et al. makes an important contribution to our understanding of how *N*-of-1 clinical trials may contribute both to individualized treatment decisions and to generalizable assessments of treatment effects in populations.

In considering whether to implement a series of *N*-of-1 trials, as compared with a more traditional design, the authors discuss the contrast between planning a meta-analysis of *N*-of-1 trials and taking advantage of a situation in which *N*-of-1 trials might be used in clinical practice to make individual treatment decisions. They note that the ''frequency of *N*-of-1 trial use in current practice is not known because individual care-focused trials comparing approved treatments and undertaken solely for personal treatment management are not typically published.'' This observation could point toward another potential tool for conducting comparative effectiveness research.

Specifically, one might ask the question, ''Why *not* publish the combined results of a collection of *N*-of-1 trials undertaken for personal treatment management?'' Working out the details of how one would operationalize such efforts will need attention. For example, avoiding publication bias that would be introduced by failing to publish equivocal or negative results and publishing only positive results; could there be a repository of such results that gets populated using a standardized reporting form from an *N*-of-1 study (perhaps even populated directly from the electronic health record [EHR]). There would be challenges, but the idea of incorporating *N*-of-1 trials more routinely into clinical practice seems promising. For medical conditions meeting the criteria outlined by the authors (a chronic and stable condition, with an intervention aimed at symptomatic improvement, rather than a permanent change in disease status and an intervention with appropriate kinetics to limit possible carryover and period effects), a more scientific and more orderly approach to individualizing patient care could be implemented on a wider scale, leading to better outcomes for individual patients, and accompanied by generalizable findings for populations. Funding for these trials, when two approved therapies are being compared, might come from health care insurance or other organizations that reimburse medical care, in lieu of imposing algorithms based entirely on ''stepped care'' (which often implies using the least expensive therapy first, followed by more expensive therapies only when the initial therapy fails).

The advisability of embedding such a research program into clinical practice depends on the particular clinical situation. For a generic therapy that is effective in 90% of people, randomization to the initial choice of therapy is less appealing than if the generic has only a 30% response rate. The benefit to individual patients would flow from a rigorous unbiased approach to finding the best therapy for that individual. One often sees the comment that two therapies may each have a 30% response rate, but that it may not be the same 30% who respond (or would respond) to both therapies. Randomization at least allows for the possibility

DOI of original article: 10.1016/j.jclinepi.2010.04.020.

The opinions expressed here are those of the author and do not necessarily represent those of Johnson & Johnson Pharmaceutical Research and Development.

* Corresponding author. Tel.: +609-730-3006; fax: +609-730-7927.
E-mail address: jberlin@its.jnj.com

0895-4356/$ - see front matter
doi: 10.1016/j.jclinepi.2010.05.006

What is new?

It may be possible to focus on individual treatment decisions and thereby gain some practical benefits to study conduct, for example, through better enrollment motivated specifically by that focus, by integrating clinical research into routine clinical care. That integration has to be acceptable to clinicians, and will no doubt require novel approaches to data collection to reduce interference with the process of care, but the potential benefits to the patients in the studies, as well as to future patients through the generation of generalizable knowledge, seem large enough to be worth the effort.

that a patient who will respond only to the usual second choice will receive that therapy first and respond sooner in the course of treatment. This could introduce a bit of a dilemma (addressed further below) in conducting trials, if it becomes difficult to keep a patient enrolled in an *N*-of-1 trial, when the first therapy provides the desired symptomatic relief.

An additional consideration in conducting *N*-of-1 studies in clinical practice is the potential to take advantage of existing electronic medical record (EMR) systems to reduce the costs of data collection. Most EMR systems are designed to record information that can help direct individual patient care. Most systems also tend not to capture specific symptomatic effectiveness data, for example, the Fibromyalgia Impact Questionnaire (FIQ) used in the trial described by Ross et al. Potentially, though, the FIQ (or other specific symptomatic data) could be captured during the routine clinical encounter (e.g., by a nurse practitioner), or by means of a handheld device provided to the patient, and then integrated into the EMR system at a relatively low cost as compared with building a separate clinical trial infrastructure. Some creativity in approaches may be needed, but the benefits of such creativity could be substantial.

Beyond these potential strategic implications of the work by Ross et al., there are a several technical points that need emphasis. The authors primarily frame the discussion of ''more periods'' vs. ''more patients'' in terms of efficiency. When there is large variability in responses within an individual (for a given treatment), adding more periods contributes relatively more information. This relative efficiency changes, however, as the balance changes between the among- and within-patient variability. The authors note, specifically, that when the within-patient variance is small, that is, the response to a particular therapy is consistent over time, additional measurements on individual patients add little because the variability in response is occurring among, rather than within, patients. This is an important concept, but there are other factors to consider, such as subject ''burden,'' and possible loss to follow-up of subjects in the *N*-of-1 studies. In the example the authors present, the study began with 58 *N*-of-1 trials but only 46 subjects completed more than one period and only 34 completed all the periods. That attrition may suggest something about the incremental value (or lack thereof) of adding more periods if, in the end, there is a loss of precision (or even validity) because of discontinuation in the study. Of note, the estimated treatment effects from the *N* of 34 analyses are generally larger than those from the *N* of 46 analyses. As the authors acknowledge, this could reflect a bias introduced by using the ''completer'' analysis. They also point out that lack of effectiveness was the main reason for dropping out of the study, supporting the concern about lack of validity. (This is the converse of the question raised above of how to keep people enrolled in a study when the initial therapy is highly effective.)

From a statistical perspective, a key finding by Zucker et al. is that models with fixed treatment effects and common variances seem to provide the most robust approach. This means that, despite variability in responses to the ''control'' therapy, the *difference* between therapies is treated as constant across patients. That is, the average benefit is assumed to apply to all patients. Whether this methodological finding generalizes to other clinical settings remains to be confirmed, but it is a crucial point that is not just part of an esoteric discussion among statisticians. It seems likely that physicians generally assume that there is patient-to-patient variability in response to a particular therapy, relative to another therapy. The message emerging from this article is that larger numbers of study subjects, or of *N*-of-1 trials, are needed to identify patient-level factors that may influence treatment response. If our interest is ultimately in being able to *predict* which patients are likely to respond to a particular therapy, this is an important paradox of which we need to be aware.

In summary, echoing the authors' points, it may be possible to focus on individual treatment decisions and thereby gain some practical benefits to study conduct, for example, through better enrollment motivated specifically by that focus, by integrating clinical research into routine clinical care. That integration has to be acceptable to clinicians, and will no doubt require novel approaches to data collection to reduce interference with the process of care, but the potential benefits to the patients in the studies, as well as to future patients through the generation of generalizable knowledge, seem large enough to be worth the effort. Importantly, Zucker et al. have taken us a step forward by addressing some key technical issues, toward making meta-analytic methods more readily available to summarize the results of *N*-of-1 trials, and toward making the impact of these studies larger.