The NEW ENGLAND JOURNAL of MEDICINE

Perspective

MAY 5, 2011

STATISTICS IN MEDICINE

Pragmatic Trials — Guides to Better Patient Care?

James H. Ware, Ph.D., and Mary Beth Hamel, M.D., M.P.H.

Related article, p. 1695

Although randomized clinical trials provide essential, high-quality evidence about the benefits and harms of medical interventions, many such trials have limited relevance to clinical practice.

The investigations are often framed in ways that fail to address patients' and clinicians' actual questions about a given treatment. For example, placebo-controlled trials of a new migraine medication help to establish its efficacy, but they may not help clinicians and patients choose between the new medication and other available treatments. Moreover, since most randomized clinical trials are efficacy trials, researchers enroll a homogeneous patient population, define treatment regimens carefully and require that they be followed assiduously, and inform neither patients nor study personnel about treatment assignments. Thus, although these trials are conducted in clinical settings, their enrolled populations and management approach don't reflect the complexity and diversity of actual clinical practice. Because of concerns about real-world applicability and about improving the quality and value of health care, "pragmatic" or "practical" trials are attracting increasing attention.[1]

Pragmatic trials are designed and conducted to answer important questions facing patients, clinicians, and policymakers.[2] They compare two or more medical interventions that are directly relevant to clinical care or health care delivery and strive to assess those interventions' effectiveness in real-world practice. They use broad eligibility criteria and recruit patients from a variety of practice settings to ensure the inclusion of the type of patients whose care will actually be influenced by the trial's results. The medical management in pragmatic trials is consistent with usual clinical care — which often means omitting study procedures such as blinding that alter the "ecology" of care. Ideally, these trials measure all the outcomes that are important to patients and decision makers, including survival, functional status, quality of life, and costs. And the duration of treatment and follow-up should be sufficient to adequately assess the treatments' benefits and risks.

In this issue of the *Journal*, Price et al. (see pages 1695–1707) report the results of two clinical

trials that were designed to address important questions about the management of asthma. In the first trial, they investigated whether leukotriene antagonists are equivalent to inhaled glucocorticoids for first-line treatment. In the second, they examined whether leukotriene antagonists are equivalent to long-acting beta-agonists as add-on therapy to inhaled glucocorticoids for patients whose asthma is poorly controlled. The trials were sponsored by Britain's Health Technology Assessment Programme for the explicit purpose of informing guidelines for asthma management in the British National Health Service.

Pragmatic trials are designed to study real-world practice and therefore represent less-perfect experiments than efficacy trials; they sacrifice internal validity to achieve generalizability.

The trials had broad eligibility criteria and were conducted at 53 primary care practices. The primary outcome was patient self-assessment by the validated Mini Asthma Quality of Life Questionnaire (Mini-AQLQ), and the trials sought to demonstrate equivalence of the treatment strategies. Since the researchers aimed to preserve the ecology of clinical care, patients and physicians were not blinded to the treatment assignment. In both studies, the criterion for equivalence was met at 2 months but not at 2 years, the primary outcome.

These trials had many strengths. They evaluated a 2-year treatment period to capture longer-term effects of treatment, and their rates of retention of subjects were impressively high. However, the investigators encountered many of the challenges characteristic of pragmatic trials. These challenges and the resulting limitations of pragmatic trials must be considered in interpreting the results.

If patients in a pragmatic trial are allowed to change or discontinue treatment readily, nonadherence can be a substantial problem. To the extent that changes in individual patients' treatment reflect normal clinical practice, an intention-to-treat analysis will provide a valid comparison of treatment *strategies*. However, nonadherence becomes a more complicated issue in a noninferiority or equivalence trial, in which it can create a bias toward a finding of equivalence.[3] Substantial crossover that results in greater similarity between the treatment regimens that are ultimately followed in two study groups will tend to yield similar outcomes in those groups. Thus, a pragmatic equivalence trial with a substantial rate of nonadherence may not demonstrate equivalence robustly. In the trial evaluating first-line asthma therapy, adherence rates, as measured by prescriptions issued, were low — 65% in the leukotriene-agonist group and only 41% in the inhaled-glucocorticoid group. Thus, the two treatment strategies may not reflect the maximal effect of the two treatments. In the add-on therapy trial, rates of adherence were only modestly better, 74% and 46%.

High rates of loss to follow-up can have a devastating effect on a pragmatic trial. Price and colleagues achieved outstanding levels of patient retention. In the two trials, 92% and 97% of patients, respectively, were included in the intention-to-treat analysis, and at least 95% of patients had a final assessment at the end of the study.

Pragmatic trials often require large sample sizes to detect small treatment effects in a heterogeneous population. Moreover, a larger study is needed to detect differences between two active treatments than between an active treatment and placebo. Price et al. enrolled a total of 658 patients in the two trials and designed the studies to have sufficient power to rule out a minimally important difference on the asthma quality-of-life scale. The availability of a validated survey instrument that was sensitive to changes in asthma symptoms contributed to their ability to address their research question without an even larger number of patients.

Unblinded treatment and clinical assessment can be an important aspect of efforts to preserve the ecology of care. In these asthma trials comparing oral and inhaled treatment, blinding would have had a significant effect on patients' experience. Moreover, patient self-assessments such as the Mini-AQLQ may be the most relevant measures of treatment effectiveness in pragmatic trials. Nevertheless, the combination of unblinded treatment and patient self-assessment undermines an important element of efficacy trials, creating a po-

tential for bias: patients' expectations about each treatment's effectiveness may influence their reporting of their quality of life. There are few if any strategies available to assess whether bias is present in a given trial. Pragmatic trials are stronger when they include both objective outcome measures (e.g., survival, test results) and subjective measures (e.g., quality-of-life surveys); concern about bias is diminished when findings from objective and subjective measures are consistent.

What, then, can we conclude from these trials? Despite the trials' limitations, the data provide encouraging evidence of equivalence or near-equivalence of the treatment strategies over the course of 2 years. Though imperfect, they provide helpful guidance for clinical care.

Yet these trials illustrate the difficulties typical of pragmatic trials. When interpreting the findings of a pragmatic trial, clinicians might ask both whether the results are valid and whether the findings are generalizable to their own patients and therefore relevant to the decisions they actually face. Pragmatic trials are designed to study real-world practice and therefore represent less-perfect experiments than efficacy trials; they sacrifice internal validity to achieve generalizability. The challenge is to keep the balance right so that the findings are likely to be both correct and applicable to clinical practice or health care delivery. It's important to recognize that, when relevant clinical trials are not available, clinicians and policymakers must turn to observational research or expert opinion to develop guidelines for care. Thus, pragmatic trials have the potential to be an important source of information to guide clinical practice and health care delivery. Careful consideration of their advantages and disadvantages and of potential strategies for mitigating their limitations will be important in improving their design and evaluation as their use is expanded.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

Dr. Ware is a professor of biostatistics at the Harvard School of Public Health, Boston, and a statistical consultant to the *Journal*.

1. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. J Chronic Dis 1967;20:637-48.
2. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. JAMA 2003;290:1624-32.
3. D'Agostino RB Sr, Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues — the encounters of academic consultants in statistics. Stat Med 2003;22:169-86.

Listening to Provenge — What a Costly Cancer Treatment Says about Future Medicare Policy

James D. Chambers, M.Pharm., M.Sc., and Peter J. Neumann, Sc.D.

In April 2010, the Food and Drug Administration (FDA) approved sipuleucel-T (Provenge), a novel cellular immunotherapy for the treatment of asymptomatic or minimally symptomatic, metastatic, castration-resistant (hormone-refractory) prostate cancer.[1,2] The pivotal clinical trial demonstrated the benefits of sipuleucel-T: an increase in median survival of 4.1 months as compared with placebo and fewer side effects than occur with docetaxel.[2] Priced at $31,000 per treatment, with a usual course of three treatments, sipuleucel-T is one of the most expensive cancer therapies ever to hit the marketplace.

In June 2010, in response to questions about whether and how its regional contractors would pay for the treatment, the Centers for Medicare and Medicaid Services (CMS) opened a national coverage analysis for sipuleucel-T. On March 30, the CMS issued a proposed decision stating that it will allow national coverage in line with the FDA-approved indication. In this decision, the CMS did not deny or restrict coverage in the off-label population but left the matter to regional contractors.[3] A final decision is expected on June 30. (The CMS has been paying for sipuleucel-T since it was approved by the FDA.)

The CMS's decision to conduct its own evaluation of sipuleucel-T raises the question of why a second government agency must review the same medical technology (with the same evidence at hand) after one government agency has already approved it. The second review risks duplication of effort, impeded access for patients, and costly delays for manufacturers seeking to commercialize innovations.

Coordination between the agencies could improve matters.