While advances in medical science have led to continued improvements in medical care and health outcomes, evidence of the comparative effectiveness of alternative management options remains inadequate for informed medical care and health policy decision making. The result is frequently suboptimal and inefficient care as well as unsustainable costs. To enhance or at least maintain quality of care as health reform and cost containment occurs, better evidence of comparative clinical and cost-effectiveness is required (1).

The American Recovery and Reinvestment Act of 2009 allocated a $1.1 billion “down payment” to support comparative effectiveness research (CER) (2). Although comparative effectiveness can be informed by synthesis of existing clinical information (systematic reviews, meta-analysis, and decision modeling) and analysis of observational data (administrative claims, electronic medical records, registries and other clinical cohorts, and case-control studies), randomized clinical trials (RCTs) are the most rigorous method of generating comparative effectiveness evidence and will necessarily occupy a central role in an expanded national CER agenda.

However, as currently designed and conducted, many RCTs are ill suited to meet the evidentiary needs implicit in the IOM definition of CER: comparison of effective interventions among patients in typical patient care settings, with decisions tailored to individual patient needs (3). Without major changes in how we conceive, design, conduct, and analyze RCTs, the nation risks spending large sums of money inefficiently to answer the wrong questions—or the right questions too late.

This article addresses several fundamental limitations of traditional RCTs for meeting CER objectives and offers 3 potentially transformational approaches to enhance their operational efficiency, analytical efficiency, and generalizability for CER.

### Enhancing Structural and Operational Efficiency

As currently conducted, RCTs are inefficient and have become more complex, time consuming, and expensive. More than 90% of industry-sponsored clinical trials experience delayed enrollment (4). In a study comparing 28 industry-sponsored trials started between 1999 and 2002 with 29 trials started between 2003 and 2006, the time from protocol approval to database lock increased by a median of 70% (4).

Several organizations have sought to streamline study start-up. In response to an analysis in Cancer and Leukemia Group B that found a median of 580 days from concept approval to phase 3 study activation (5), the National Cancer Institute established an operational efficiency working group to reduce study activation time by at least 50%, increase the proportion of studies reaching accrual targets, and improve timely study completion (6).

The National Institutes of Health’s Clinical and Translational Science Award recipients are documenting study start-up metrics as a first step to fostering improvements (7). The National Cancer Institute, the CEO Roundtable, Cancer Centers, and Cooperative Groups developed standard terms for clinical trial agreements as a starting point for negotiations between study sponsors and clinical sites (8). The Institute of Medicine’s Drug Forum also commissioned development of a template clinical research agreement (9).

Through its Critical Path Program, the U.S. Food and Drug Administration (FDA) established the Clinical Trials Transformation Initiative (CTTI), a public–private partnership whose goal is to improve the quality and efficiency of clinical trials (10). The CTTI is hosted by Duke University and has broad representation from more than 50 member organizations, including academia, government, industry, clinical investigators, and patient advocates (11). The CTTI works by generating empirical data on how clinical trials are currently conducted and how they may be improved. Initial priorities for study include design principles, data quality and quantity (including monitoring), study start-up, and adverse event reporting.

One of CTTI’s projects is addressing site monitoring, an area that has been estimated to absorb 25% to 30% of phase 3 trial costs (12) and for which there is widespread agreement that improved efficiency is needed. The CTTI is determining the current range of monitoring practices for RCTs used by the National Institutes of Health, academic
institutions, and industry; assessing the quality objectives of monitoring; and determining the performance of various monitoring practices in meeting quality objectives. This project will provide criteria to help sponsors select the most appropriate monitoring methods for a trial, thereby improving quality while optimizing resources.

Collectively, these efforts are generating empirical evidence and developing the mechanisms to improve clinical trial efficiency. In conjunction with other improvements, including those described below, the resulting changes in clinical trial practices will increase the feasibility of mounting the scale and scope of RCTs required to evaluate the comparative effectiveness of medical care.

**Analytical Efficiency: The Potential Role of Bayesian and Adaptive Approaches**

The traditional frequentist school has provided a solid foundation for medical statistics. But the artificial division of results into “significant” and “nonsignificant” is better suited for one-time dichotomous decisions, such as regulatory approval, and is not the best model for comparing interventions as evidence accumulates over time, as occurs in a dynamic medical care system.

With traditional trials and analytical methods, it is difficult to make optimal use of relevant existing, ancillary, or new evidence as it arises during a trial, and thus such methods often are not well suited to facilitate clinical and policy decision making. Furthermore, real-world CER can be “noisier” than a standard RCT. Standard statistical techniques require increased sample sizes, in part because of the resulting additional variability and in part when trials compare several active treatments whose effectiveness differs by relatively small amounts.

Designs that use features that change or “adapt” in response to information generated during the trial can be more efficient than standard approaches. Although many standard RCTs are adaptive in limited ways (for example, those with interim monitoring and stopping rules), the frequentist paradigm inhibits adaptation because of the requirement to prespecify all possible study outcomes, which in turn requires some rigidity in design. The Bayesian approach, using formal, probabilistic statements of uncertainty based on the combination of all sources of information, provides a more flexible tool for clinical and policy decision making. The Bayesian approach increases the feasibility of mounting the scale and scope of RCTs required to evaluate the comparative effectiveness of medical care.

Bayesian and adaptive analytical approaches can allow these interventions to be added and less effective ones dropped without restarting the trial; therefore, at any given time, the trial is comparing the alternatives most relevant to current clinical practice. This dynamic “learning adaptive” feature (analogous to the Institute of Medicine Evidence-Based Medicine Roundtable’s “learning health care system” [14]) improves both the timeliness and clinical relevance of trial results.

The following example shows how this model operates. A standard comparative effectiveness trial design of 4 alternative strategies for HIV infection treatment starts with the hypothesis of equal effectiveness of all 4 treatments. In contrast, as the trial progresses, the Bayesian approach answers the pragmatic questions: “What is the probability that the favored therapy is the best of the 4 therapies?” and “What is the probability that the currently worst therapy will turn out to be best?” (15). If this latter probability is low enough, the trialists can drop that treatment even if it is not, by conventional statistical testing, worse than other treatments. Newly developed HIV treatment strategies also can enter the trial, thus focusing patient resources on the most relevant treatment comparison. Bayesian and adaptive designs are particularly useful for rapidly evolving interventions (such as devices, procedures, practices, and systems interventions), especially when outcomes occur soon enough to permit adaptation of the trial design. They should also prove useful for clinical studies generated by such conditional coverage schemes as Medicare’s Coverage with Evidence Development policy by adding onto an existing evidence base and “adapting” studies into community care settings of interest to payers and patients (16, 17).

Random allocation need not be equal between trial arms or patient subgroups. Probabilities of each intervention being the best can be updated and random allocation probabilities revised, so that more patients are allocated to the most promising strategies as evidence accumulates. This flexibility can also permit Bayesian trials to focus experimentation on clinically relevant subgroups, which could facilitate tailoring strategies to particular patients, a key element of CER.

Experience with Bayesian adaptive approaches has been growing in recent years. Early-phase cancer trials are commonly performed using Bayesian designs (18). In 2005, the FDA released a draft guidance document for the use of Bayesian methods in device trials (19), and the FDA Center for Drug Evaluation and Research and Center for Biologics Evaluation and Research have accepted Bayesian trial designs. The I-Spy 2-design trial (20) is a particularly interesting implementation of this new paradigm, being less a single trial than a set of ongoing evaluation processes that simultaneously evaluate treatments and biomarkers in women with advanced-stage breast cancer. Nearly all large pharmaceutical companies and many biotechnology and device companies have implemented Bayesian adaptive designs in their product development process (21, 22).

Bayesian and adaptive approaches to RCT design and analysis are among several promising novel design and an-
alytic approaches that could help meet CER evidence challenges. These challenges include the need to compare multiple active treatment strategies in real-world settings, to focus experimental resources on the most promising approaches, to identify patient subgroups in which treatments are more (or less) effective, to introduce new treatments into the evaluation process as quickly as possible, and to make optimal use of all existing experimental information when a study is designed and as it is conducted. If the promised potential of CER is to be realized, we will need to keep exploring the possibilities for new methods for statistical trial design.

**Pragmatic Clinical Trials: RCTs Designed for Decision Makers**

A defining objective of CER is to provide information to help patients, consumers, clinicians, and payers make more informed clinical and health policy decisions. However, many RCTs exclude clinically relevant patient subgroups (as defined by age, sex, race, ethnicity, and comorbid conditions), commonly used comparator interventions, important patient outcomes (such as quality of life and longer-term effect), and nonexpert providers (23). These exclusions diminish the relevance of the trial results to some important clinical and policy decisions.

Although exclusions of clinically important subgroups are sometimes due to risk–benefit concerns, they also reflect the purpose of most RCTs: to determine an intervention’s net benefit under ideal circumstances (efficacy), either to satisfy FDA marketing approval requirements or to provide insights into disease etiology and underlying mechanisms of disease (24). These goals lead to tightly controlled study designs that are consequently less likely to reflect the conditions under which interventions are used in common clinical practice. The resulting trials frequently do not reach their potential value for health care decision making, which is a serious waste of resources.

The RCTs whose explicit purpose is to be most informative to decision makers are called pragmatic or practical clinical trials (PCTs) (23–26) and are well aligned with the purpose of CER. Common elements of such trials include clinically effective comparators, study patients with common comorbid conditions and diverse demographic characteristics, and providers from community settings. Primary and secondary outcomes are patient-centered, chosen to reflect what matters most to patients and clinicians.

The Pragmatic-Explanatory Continuum Indicator Summary (PRECIS) provides a useful framework to help researchers design trials that inform health care decisions (23). This tool identifies important domains of trial design (such as eligibility criteria, patient adherence, and practitioner expertise) that should be considered during PCT protocol development. Strategies to involve practicing clinicians, payers, and particularly patients and consumers in clinical trial design are not well developed or widely practiced, and these will have to be further developed to maximize the value of CER (3).

The very aspects of PCTs that make them most useful to decision makers sometimes make them more difficult to interpret from an explanatory perspective. It may not be clear from a PCT how biological, etiologic, or behavioral mechanisms interact to produce the observed clinical outcomes. For example, if efficacy has been established in a narrow population but effectiveness is not seen in a broader population, is the result of misapplication of the intervention (poor patient adherence or provider performance) or the absence of an efficacy in a broader patient group? Addressing this important question may require RCT analytic approaches that borrow from the methods of observational studies, for which bias and confounding are endemic problems.

**Discussion**

We identify 3 key issues—operational efficiency, analytical efficiency, and pragmatic approaches to trial design—that, if embraced, will advance the CER initiative by more efficiently generating valid, generalizable evidence from randomized trials.

Operational efficiency of trials is particularly germane to the needs of CER because comparisons of multiple effective treatments will require larger sample sizes to reliably detect differences and patients, clinicians, and payers will demand comparative information in shorter time frames. Comparative effectiveness research will benefit from the activities of several ongoing public–private initiatives seeking to improve RCT efficiency while maintaining quality and making them more informative for patients, clinicians, and payers.

Comparative effectiveness research trials must be designed for the dynamic, unique needs of community medical practice. Bayesian and adaptive trial methods are not new, but they have not yet been applied to CER. Applying these methods to CER is consistent with the concept of the “learning healthcare system” in that they allow flexible, adaptive, cumulative learning to be incorporated during the conduct of trials. The ultimate result should be clinically relevant, timely information to inform clinical and policy decisions. These methods may be especially useful for rapidly evolving interventions, patient or subgroup characteristics that predict response to alternative management strategies, and conditional coverage schemes.

The PCT is more concept than method, its form being based on the requirements of the clinical and policy questions it is designed to address. Properly designed PCTs require meaningful involvement of patients, physicians, payers, health policymakers, and other relevant stakeholders to ensure that the research will meet their objectives and decision-making needs.

In the U.S. medical system, we face countless questions about the relative benefits and risks of both new and existing interventions in the setting of usual care. As the CER initiative begins to tackle this challenge by prioritiz-
ing the questions and recommending methods to address them, there will be limits to the value of observational data alone to inform these decisions.

Randomized clinical trials offer reliable information and must have a prominent place in the CER agenda. The improved approaches to RCTs that we have discussed, by requiring a thorough CER-focused rethinking of RCT study design principles, operational procedures, statistical methods and rules by academic disciplines, funding agencies, regulatory authorities, and business interests that comprise the clinical trial research enterprise, will ensure the feasibility of using this methodology with greater efficiency, generalizability, and responsiveness to the changing health care system.

From United BioSource Corporation, Bethesda, and Johns Hopkins Schools of Medicine and Public Health and Center for Medical Technology Policy, Baltimore, Maryland; Whartson School and Leonard Davis Institute of Health Economics, University of Pennsylvania, and School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; Duke Translational Medicine Institute, Duke University, Durham, North Carolina; and Berry Consultants and University of Central Florida College of Medicine, Orlando, Florida.

Potential Financial Conflicts of Interest: Dr. Luce: Director, PACE (Pragmatic Approaches to Comparative Effectiveness) Initiative. PACE activities include evaluating and fostering novel CER trial methodologies, including Bayesian adaptive methods. PACE receives unrestricted funding from private organizations, including life sciences manufacturers.


Dr. Kramer: Executive Director of CTTI.

Dr. Tunis: Member, Institute of Medicine Committee on Initial Priorities for Comparative Effectiveness Research.

Dr. Tunis and Ms. Whicher: Director and Research Associate, respectively, for the Center for Medical Technology Policy, a nonprofit organization that promotes the development of methods for CER, including PCTs.

The authors received no compensation for writing this manuscript.

Requests for Single Reprints: Bryan R. Luce, PhD, MBA, United BioSource Corporation, 7101 Wisconsin Avenue, Suite 600, Bethesda, MD 20814; e-mail, bryan.luce@unitedbiosource.com.

Current author addresses are available at www.annals.org.

References


Current Author Addresses: Dr. Luce: United BioSource Corporation, 7101 Wisconsin Avenue, Suite 600, Bethesda, MD 20814.
Dr. Kramer: PO Box 17969, Durham, NC 27715.
Dr. Goodman: Department of Oncology, Johns Hopkins School of Medicine, 550 North Broadway, Suite 1103, Baltimore, MD 21205.
Dr. Connor: Berry Consultants, 2534 Lake Debra Drive, #108, Orlando, FL 32825.

Dr. Tunis: Center for Medical Technology Policy, Inner Harbor Center, 400 East Pratt Street, Suite 808, Baltimore, MD 21202.
Ms. Whicher: Center for Medical Technology Policy, Inner Harbor Center, 400 East Pratt Street, Suite 814, Baltimore, MD 21202.
Dr. Schwartz: Department of Medicine, School of Medicine, University of Pennsylvania, 1123 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021.