

Practical Clinical Trials

Increasing the Value of Clinical Research for Decision Making in Clinical and Health Policy

Sean R. Tunis, MD, MSc

Daniel B. Stryer, MD

Carolyn M. Clancy, MD

THE NEED FOR CAREFUL SCIENTIFIC evaluation of clinical practice became a prominent focus during the second half of the 20th century. Demonstration of pervasive and persistent unexplained variability in clinical practice¹ and high rates of inappropriate care,² combined with increased expenditures, have fueled a steadily increasing demand for evidence of clinical effectiveness. The limited amount of high-quality evidence is recognized to be partly responsible for geographic variation, inappropriate care, and the limited success of quality improvement efforts.^{3,4} As a result, increased attention is being directed to the development of methods that can provide valid and reliable information about what works best in health care.

THE NEED FOR HIGH-QUALITY EVIDENCE

Among the primary audiences for higher-quality evidence are clinical and health policy decision makers, including patients, physicians, payers, purchasers, health care administrators, and public health policymakers. Patients and physicians increasingly seek to combine their personal beliefs about health care choices with attention to high-quality evidence in making individual decisions about care. Medical professional societies produce guidelines to assist physicians and

Decision makers in health care are increasingly interested in using high-quality scientific evidence to support clinical and health policy choices; however, the quality of available scientific evidence is often found to be inadequate. Reliable evidence is essential to improve health care quality and to support efficient use of limited resources. The widespread gaps in evidence-based knowledge suggest that systematic flaws exist in the production of scientific evidence, in part because there is no consistent effort to conduct clinical trials designed to meet the needs of decision makers. Clinical trials for which the hypothesis and study design are developed specifically to answer the questions faced by decision makers are called pragmatic or practical clinical trials (PCTs). The characteristic features of PCTs are that they (1) select clinically relevant alternative interventions to compare, (2) include a diverse population of study participants, (3) recruit participants from heterogeneous practice settings, and (4) collect data on a broad range of health outcomes. The supply of PCTs is limited primarily because the major funders of clinical research, the National Institutes of Health and the medical products industry, do not focus on supporting such trials. Increasing the supply of PCTs will depend on the development of a mechanism to establish priorities for these studies, significant expansion of an infrastructure to conduct clinical research within the health care delivery system, more reliance on high-quality evidence by health care decision makers, and a substantial increase in public and private funding for these studies. For these changes to occur, clinical and health policy decision makers will need to become more involved in all aspects of clinical research, including priority setting, infrastructure development, and funding.

JAMA. 2003;290:1624-1632

www.jama.com

patients in making medical decisions.⁵ Health insurers and managed care organizations increasingly depend on systematic reviews and technology assessments to support quality improvement efforts and to develop coverage and payment policy.^{6,7} Hospitals and health sys-

Author Affiliations: Centers for Medicare & Medicaid Services, Baltimore, Md (Dr Tunis); and Agency for Healthcare Research and Quality, Rockville, Md (Drs Stryer and Clancy).

Corresponding Author and Reprints: Sean R. Tunis, MD, MSc, Office of Clinical Standards and Quality, Centers for Medicare & Medicaid Services, 7500 Security Blvd, Mailstop S3-02-01, Baltimore, MD 21244 (e-mail: stunis@cms.hhs.gov).

tems use these documents to make decisions about capital investments such as equipment purchasing and construction (BOX).

The current clinical research enterprise in the United States is not consistently producing an adequate supply of information to meet the needs of clinical and health policy decision makers.⁸⁻¹⁰ The inability to address many common, important clinical questions, despite a significant increase in public and private funding for clinical research, suggests a systemic problem in the production of clinical research. This article explains the impact of knowledge gaps on health care decision makers, describes the features of clinical trials that would more reliably answer the practical questions they face, and discusses why the current clinical research enterprise fails to address many important practical questions. This article also proposes strategies to address the current shortage of clinical trials to meet these needs.

PREVALENCE AND IMPACT OF KNOWLEDGE GAPS

The prevalence and significance of gaps in knowledge about clinical effectiveness are most readily appreciated by reviewing the results of most systematic literature reviews, technology assessments, and clinical practice guidelines. These reports are generally produced to provide comprehensive reliable information for decision makers and usually address common conditions with large aggregate cost, morbidity, and public health importance. A consistent finding of these reviews is that the quality of evidence available to answer the critical questions identified by experts is suboptimal. For example, a systematic review of newer pharmacologic agents for depression concludes that few studies provided data on the long-term effectiveness of treatment, the functional status of patients, or the outcomes of patients treated in typical practice settings.¹¹ Furthermore, few studies compared the older inexpensive agents with newer agents in terms of adverse effects and

Box. Uses of Evidence in Decision Making

Physician/Patient Decision Making

Of existing diagnostic or treatment alternatives, which makes the most sense for an individual patient?

Choosing Plans or Physicians

Which plan or physician is likely to provide high-quality care?

Practice Guidelines

What is the best approach for patients with selected conditions?

Quality Measurement and Improvement

How can evidence-based clinical performance be assessed? Do improvement programs result in enhanced clinical care?

Product Purchasing and Formulary Selection

How does this product compare with existing alternatives?

Benefit and Coverage Decisions

Should a new service be reimbursed and for which patients?

Organizational and Management Decisions

Does a hospitalist program decrease costs and improve outcomes?

Program Financing and Priority Setting

Which services represent the best value for additional investments?

Product Approval

Should this product be approved and, if so, for which indications?

clinical efficacy. Most well-done systematic reviews and clinical guidelines reach similar conclusions about the quality of evidence associated with common clinical problems.

These gaps in evidence undermine efforts to improve the scientific basis of health care decisions in several ways. Organizations that develop evidence-based clinical practice guidelines may not be able to develop clear, specific recommendations.¹² For example, the background report for clinical guidelines on outpatient management of exacerbations of chronic obstructive pulmonary disease found that although numerous industry-sponsored clinical trials reported minor differences in the antimicrobial activity of alternative broad-spectrum antibiotics, no trials had been performed to determine whether any of the newer broad-spectrum antibiotics were better than older generic antibiotics or even placebo (for mild exacerbation).^{13,14} As a result, the guideline could not provide definitive recommendations on the appropriate choice of anti-

biotics for chronic obstructive pulmonary disease exacerbations.

The limited quantity and quality of available scientific information also impede the efforts of public and private health insurers in developing evidence-based coverage policies for many new and existing technologies.^{7,15} Poor-quality studies of new technologies can lead to millions of dollars being allocated for new technologies for which the long-term benefits and risks have not been determined.⁴ Minimally invasive technologies for treatment of benign prostatic hyperplasia (BPH) are in widespread use, yet no clinical trials have been performed to compare the risks and benefits of these treatments with standard surgical interventions.¹⁶ The Medicare program has spent millions of dollars per year for home use of special beds for patients with pressure ulcers, despite the fact that no well-designed study demonstrates that they improve healing of these ulcers.¹⁷ The limited production of this body of research becomes increasingly problematic as major increases in

public funding for basic research generate an expanding range of potentially valuable technologies in need of careful objective evaluation.

DESIGNING CLINICAL RESEARCH FOR DECISION MAKERS

Health services research and outcomes research have made important contributions toward the effective translation of clinical research discoveries to clinical practice and health policy.¹⁸⁻²⁰ However, observational and other nonexperimental methods may not provide sufficiently robust information regarding the comparative effectiveness of alternative clinical interventions, primarily because of their high susceptibility to selection bias and confounding.^{21,22} The recent findings demonstrating the adverse effects and lack of effectiveness of hormone therapy despite consistent findings of benefit from observational studies highlight the limitations of nonexperimental studies.²³

Clinical trials designed to assist health care decision makers, referred to as *pragmatic clinical trials* or *practical clinical trials* (PCTs), are defined as trials for which the hypothesis and study design are formulated based on information needed to make a decision.²⁴ They are distinguished from *explanatory clinical trials*, for which the goal is to better understand how and why an intervention works. Explanatory trials are designed to maximize the chance that some biological effect of a new treatment will be revealed by the study. The PCTs address practical questions about the risks, benefits, and costs of an intervention as they would occur in routine clinical practice.²⁵ The most distinctive features of PCTs are that they select clinically relevant interventions to compare, include a diverse population of study participants, recruit participants from a variety of practice settings, and collect data on a broad range of health outcomes.

Compare Clinically Relevant Alternatives

Often PCTs are designed as head-to-head comparisons of viable alternative

clinical strategies. These comparative studies have the potential to alter clinical decisions profoundly, because they are derived from practical choices facing patients and their physicians. In contrast, many of the clinical studies comparing individual agents with placebo may not provide a sound basis for choosing from among acceptable alternatives.

For example, a PCT conducted by the Department of Veterans Affairs (VA) compared the benefits of pharmacologic therapy with terazosin hydrochloride, finasteride, or both for treatment of symptoms of BPH.²⁶ Both drugs are approved by the Food and Drug Administration (FDA) for this use based on trials comparing each drug with a placebo. There is no incentive or requirement for manufacturers to initiate studies to compare the products. The VA study²⁶ randomized 1229 men with BPH and demonstrated that terazosin was more effective than finasteride for reducing BPH symptoms at 1 year. Once initiated, both manufacturers contributed to the design and funding for the study. The National Institutes of Health (NIH) funded a study that compared doxazosin mesylate, finasteride, and both drugs in 3047 patients with BPH for a mean of 4.5 years. This study revealed that the combination of the 2 drugs was significantly more effective at delaying progression of symptoms than either drug alone.²⁷

In addition, PCTs can address non-pharmacologic alternatives. Despite back pain being among the most prevalent complaints in primary care, few high-quality studies have been performed to compare results of the numerous existing treatment alternatives. A PCT of therapy for recent-onset low back pain randomized 323 patients to 1 of 3 widely used alternative adjunct treatment strategies: physical therapy, chiropractic care, and self-care (the principles of which are described in an educational booklet).²⁸ The study showed that physical therapy and chiropractic care increased patient satisfaction and marginally reduced symptoms compared with the self-care principles outlined in the booklet; however, there were no differences between

the 3 study groups in function or rates of recurrence. The educational booklet that outlined self-care was substantially less expensive.²⁸ A study of similar design demonstrated that therapeutic massage was more effective and less costly than acupuncture in treating low back pain.²⁹

Enroll a Diverse Study Population

Typically, PCTs include a more diverse study population by having broad inclusion criteria and fewer exclusion criteria when enrolling patients. The goal is to enroll patients in the trial with characteristics that reflect the range and distribution of patients observed in clinical practice for a particular problem. This approach addresses the common concern of decision makers about the applicability of results from studies with restricted eligibility criteria.^{30,31} It can also ensure that the higher-risk patients likely to have the greatest benefit from some treatment are not excluded from clinical trials.³²

Most clinical trials of antidepressants have excluded elderly patients and focus on patients with major depression rather than patients with less severe depression, which is a more prevalent problem. Williams et al³³ enrolled a substantial number of elderly patients in a PCT of 415 patients with minor depression or dysthymia, comparing the effectiveness of oral antidepressants, problem-solving treatment (a form of cognitive therapy), and placebo. The study found that treatment with paroxetine showed greater improvement in depressive symptoms and mental health function compared with problem-solving treatment or placebo. Reliable and relevant information such as this is more likely to convince physicians to provide medication for their elderly patients with this common cause of significant reversible morbidity.

Because physicians must often treat patients based on the likely rather than confirmed diagnosis, studies that enroll patients based on presenting symptoms rather than definitive test results may be of great practical value. A PCT of patients with sinusitis compared the effectiveness of antibiotics with and

without nasal corticosteroids.³⁴ Because primary care physicians typically use sinus radiographs to evaluate patients with suspected sinusitis, radiographic evidence of sinusitis was used as a study entry criterion (rather than sinus puncture). This approach maximized the generalizability of the study results by enrolling a patient population that physicians would recognize as similar to patients observed in daily practice. Another PCT randomized 478 patients to 2 alternative strategies for managing patients who presented with dyspepsia: (1) blood testing for *Helicobacter pylori* with patients with positive test results referred for endoscopy or (2) empirical treatment with acid-suppressive drugs.³⁵ Neither study group involved routine referral for endoscopy, replicating the standard of care for this common primary care problem.

Recruit From a Variety of Practice Settings

Also, PCTs may improve external validity by including a wider range of physicians and settings to which the study will be applicable. Pragmatic clinical trials have often been conducted in community-based settings, generally clinics and physician offices at which the primary activity is clinical care. Enrolling patients from a more diverse group of practice settings also allows for some variability in how the study intervention and associated clinical care will be provided.²⁴ The ancillary care that these patients receive is more likely to reflect the average care patients would receive outside the research context. In the study of nasal corticosteroids for sinusitis, 18 of the 22 study sites were community-based primary care or otolaryngology clinics. Dolor et al³⁴ noted that this selection of study settings would ensure that the study sample represented the patient population typically observed in general practice. Outcomes of risky or technically demanding procedures are particularly likely to vary across sites, making it especially valuable to have results that reflect outcomes across a range of institutions.

The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) compared treatment outcomes for various drug treatments for hypertension in more than 42 000 patients enrolled from 623 primary care clinics.³⁶ Many of the study sites had not previously been sites for clinical research. The study demonstrated that use of inexpensive diuretic medication resulted in a lower incidence of major cardiovascular events compared with the newer and more expensive classes of antihypertensive agents, including calcium channel blockers, angiotensin-converting enzyme inhibitors, and α -blockers.^{37,38} Physicians and policymakers can be confident that the reported outcomes in this study are likely to predict results that will be observed across a wide range of practice settings. This eliminates a common obstacle to physician implementation of clinical research findings.

Measure a Broad Range of Relevant Health Outcomes

Selection of the outcomes to be measured in PCTs is based on the most important anticipated effects of the intervention, taking into account those outcomes of greatest relevance to decision makers. As a result, the study end points collected in PCTs include a broad range of functional outcomes, including quality of life, symptom severity, satisfaction, and costs, as well as more traditional end points, such as mortality and major morbidity.^{39,40} The low back pain study by Cherkin et al²⁸ reported disability days, satisfaction, and ability to function. Many large trials of cardiovascular interventions now report general and disease-specific quality of life rather than focusing only on mortality or major nonfatal outcomes such as stroke or myocardial infarction.

In addition, the period of follow-up for PCTs is often longer than in traditional clinical trials to better reflect more of the natural history of a disease. Many therapies have different results in the short term than in the long term, a phenomenon commonly observed in surgical trials that usually demonstrate

high short-term risks associated with surgery.³² Two PCTs that compared the results of immediate surgical repair with surveillance for small abdominal aortic aneurysms followed up patients for 4.9 years in 1 trial and 8.0 years in the other trial.^{41,42} Previous studies on this question had been much shorter and had left substantial uncertainty about the safety of deferring surgical repair of small aneurysms. Both studies demonstrated that survival was not improved by elective surgical repair of these aneurysms, a result that could lead to the avoidance of many unnecessary surgical procedures.

Inclusion of economic outcomes in clinical trials provides information that has a significant impact on clinical and policy decisions.⁴³ A randomized study evaluated the results of treatment with fluoxetine compared with tricyclic antidepressants for the initial management of mild depression.⁴⁴ Primary outcomes of the study included the type and frequency of adverse effects, the number of patient visits associated with these adverse effects, and the rate at which patients were switched from their originally assigned medication to another agent. The study concluded that, despite the higher cost of fluoxetine, total costs of care were no different between the 2 groups because the incidence of adverse effects and frequency of physician visits were lower for patients initially treated with fluoxetine. The findings provided important input for formulary decision makers and were also given substantial weight in the development of evidence-based clinical guidelines on the management of depression.⁴⁵

WHY IS THE SUPPLY OF PCTs INADEQUATE?

The supply of PCTs is currently inadequate in large measure because trials often require substantial resources, and the current level of public- and private-sector funding for these studies is inadequate. Neither of the major sources of funding for clinical research in the United States—the NIH and the medical products industry—has as a primary mission the goal of ensuring that

studies are performed to address clinical questions important to decision makers. The funding entities that are focused on supporting these studies have relatively limited resources.

Practical clinical trials may be extremely costly to conduct because they can require large sample sizes and may follow up patients for a long period. The 7-year, 42000-patient ALLHAT study is estimated to have cost approximately \$120 million to complete (C. Furberg, MD, PhD, written communication, January 9, 2003). The National Emphysema Treatment Trial enrolled 1200 patients in the surgical and medical therapy groups, took approximately 6 years to complete, and cost more than \$35 million for the research, with substantially more than this amount paid by Medicare for the clinical care of patients in the trial (G. Weinman, MD, written communication, April 16, 2003). Even studies of relatively modest size and simple design can have substantial costs. An NIH study of St John's wort and citalopram will enroll 300 patients with mild depression and follow them up for 20 weeks, with an estimated cost of \$4 million over 4 years.⁴⁶

Although PCTs are not a major focus for the NIH, they have funded a number of important PCTs, most of which have provided crucial information for clinical and health policy decision makers. The NIH does not, however, have any organized systematic mechanism in place to identify the high-priority questions of decision makers that might need to be resolved through a PCT. Estimates of the distribution of NIH extramural grants indicate that approximately 70% of NIH grants are devoted to basic biomedical research, with the remainder going to clinical research of all types.⁴⁷ The NIH scientific review committees, as currently constituted, are more familiar with traditional explanatory trials and have relatively limited experience with PCTs. Consistent with its primary mission and funding priorities, the NIH has been successful at biomedical discovery but has been less consistent in supporting clinical research that ensures that these

discoveries will be effectively translated into improved patient care and public health.¹⁰

Some public research funding organizations have an explicit commitment to supporting research that addresses practical questions of health care decision makers. The Agency for Healthcare Research and Quality (AHRQ) has funded several important PCTs^{48,49} but devotes most of its resources to health services research, quality measurement, and patient safety. According to AHRQ (L. Levine, oral communication, February 24, 2002), total spending on PCTs was approximately \$30 million in fiscal year 2001, and the current agency budget of approximately \$300 million would not allow for a significant expansion for such studies while maintaining its other core activities. The AHRQ and FDA jointly oversee the Centers for Education and Research in Therapeutics, which were established for the purpose of improving the effective use of pharmaceuticals. The Centers for Education and Research in Therapeutics are directed to accomplish this by conducting "essential research studies on drugs and therapies that are not being performed by the pharmaceutical industry or the NIH" and communicating these findings to practicing physicians.⁵⁰ The annual funding by the Centers for Education and Research in Therapeutics of \$7 million is adequate to identify, but not to support, important PCTs related to pharmaceutical therapy.

The Cooperative Studies Program of the VA has produced numerous studies designed to support clinical and policy decisions. This focus of activity emerges directly from the purpose of that research program: to support a research agenda designed to improve the effectiveness of care for veterans. This program funded some of the earliest trials clearly demonstrating the value of coronary artery bypass graft surgery.⁵¹ More recently, the VA funded an evaluation of arthroscopic surgery for osteoarthritis of the knee, which showed that the results of placebo surgery were no different than results from actual lavage and debride-

ment.⁵² Although the clinical implications of the study are still being debated, the Department of Veterans Affairs issued an advisory to its physicians, recommending that they not perform this surgery in patients similar to those in the clinical trial.⁵³ The study also prompted the Medicare program to consider specific limitations for use of this procedure, guided primarily by results of the VA study. With approximately \$55 million to support such trials, the VA is an important source of PCTs in the United States, but the number of important questions is far beyond this level of funding.

Industry funding for clinical trials is several times greater than NIH spending (\$4.1 billion vs \$850 million in 2000)⁵⁴; however, most of the clinical research supported by the pharmaceutical and medical device industries is designed primarily to meet the regulatory requirements for marketing approval by the FDA. More than 90% of industry spending on clinical trials supports phase 1, 2, and 3 studies required for regulatory approval, with the remaining 10% spent on phase 4 clinical trials.⁵⁴ Most PCTs would be classified as phase 4 trials.

Most devices reviewed by the FDA are cleared for marketing based on the similarity of those devices to previously approved devices, and there is no requirement for clinical studies that directly evaluate the clinical effectiveness of most devices. Once FDA approved, devices may be widely used off label for clinical purposes that have never been carefully evaluated. Most of the pulmonary artery catheters in modern use have been approved by the FDA without clinical trial data based on the similarity of these catheters to those in use before 1976 (the year that the FDA began to regulate medical devices). Large randomized trials have recently shown that the pulmonary artery catheter does not provide clinical benefit compared with standard care in elderly high-risk surgical patients who require intensive care.⁵⁵ Most of the therapies in common use for treatment of pressure ulcers and other chronic wounds, such as electrical stimulation and hyperbaric oxygen, have never been

the subject of high-quality trials or received FDA approval for treatment of wounds.

For pharmaceuticals, the FDA requires placebo-controlled studies for approval and generally does not require trials that compare a new drug with existing alternatives, unless a company is specifically seeking FDA approval to make a claim of superiority to another marketed product. These comparative studies are also unattractive to drug manufacturers if there is any chance that the results might reflect poorly on their product. Therefore, industry-sponsored drug studies will usually not provide the comparative information about risks and benefits sought by most health care decision makers. Furthermore, because the FDA requires information about efficacy rather than effectiveness, the study populations enrolled in FDA-approval studies are highly selected, in contrast with the more diverse study populations that enhance the generalizability of results from PCTs.

Although medical product companies increasingly support additional clinical studies after FDA approval, postmarket research priorities for industry are understandably guided by the objective of maximizing market share rather than the goal of addressing critical questions concerning optimal clinical use.⁵⁰ For example, a longitudinal observational study was recently conducted that recruited more than 200 000 ambulatory, postmenopausal women older than 50 years from more than 4000 primary care practices.⁵⁶ The substantial cost of this study was borne by a manufacturer of a prescription drug for osteoporosis and an organization representing the industry that produces bone mineral density measurement devices. The main result of the study was that nearly half of women older than 50 years may be at increased risk of osteoporotic fractures. These results suggest that physicians should much more aggressively screen for and treat women at risk for osteoporotic fractures, findings that are clearly aligned with the financial interests of both major study sponsors. It would be unlikely that the same sponsors would de-

vote resources to the equally important question of whether lifestyle modification was as good as drug therapy in reducing the risk of fractures in these at-risk women.

STRATEGIES FOR IMPROVEMENT

We propose a set of strategies that we believe would promote more consistent production of clinical trials that meet the needs of health care decision makers. These strategies would call for sustained effort and substantial resources from all affected stakeholders.

Systematic Identification and Prioritization of Knowledge Gaps

Currently, there is no institution with the responsibility for identifying and prioritizing important unanswered clinical research questions that would be useful to clinical and health policy decision makers. It is essential that this function be established. Ideally, the responsible entity would have majority representation from the decision makers in need of this information, including representatives of payers, purchasers, medical professionals, and patients.⁵⁷ One possibility would be for the Institute of Medicine to serve this function, providing staff and supportive infrastructure within a neutral scientific environment. The list generated by this group would be available to public and private funders as they develop new initiatives and make their funding decisions. The degree to which funders supported studies from this list would highlight whether research funders were meeting important public health needs. To the extent that other health care stakeholders begin to devote additional resources to clinical research, the priority list will also serve as a guide to research most likely to generate reliable usable information.

There are a number of sources from which high-priority questions for PCTs could be identified. Virtually every clinical guideline, technology assessment, systematic review, and consensus report includes a section that lists specific clinical research priorities. These priorities de-

serve special attention because of the systematic and comprehensive method by which they have been generated.

Decision Makers Insist on High-Quality Evidence in Making Decisions

The production of high-quality clinical trials will increase significantly when health care decision makers decide to consistently base their decisions on high-quality evidence. Research sponsors (public and private) will be motivated to provide the type of clinical research required by decision makers. Payers and purchasers can clearly indicate to the drug and device industry that favorable coverage and payment decisions will be expedited by reliable evidence from PCTs. In particular, manufacturers will be motivated to perform head-to-head comparative trials if these are required to justify payments higher than the existing less expensive alternatives. Physicians and medical professional organizations can also increase the degree to which care of individual patients and professional society clinical policy are guided by attention to reliable evidence.

Patients and their advocacy organizations have a critical stake in the quality of clinical research. The fundamental premise of improving health care through informed patient choice cannot be realized if unreliable evidence is used to inform patients. Patient advocacy organizations should become more insistent on the production of high-quality evidence and work with research funders to increase support for PCTs. These groups could also facilitate the research by encouraging patients, physicians, and health care organizations to participate in clinical trials in those situations for which current evidence is uncertain. Many years and lives could have been saved had advocates worked to ensure the rapid completion of clinical trials on bone marrow transplantation for breast cancer, instead of putting pressure on state and federal policymakers to force payers to cover this procedure when supportive high-quality data were lacking.⁵⁸

Create Operational Infrastructure

To conduct PCTs efficiently, a durable infrastructure within the primary care setting is needed. Considerable effort and resources will be required to develop additional networks of physicians willing and able to participate in clinical research. The HMO Research Network, a consortium of large health maintenance organizations, including Kaiser Northern California, Group Health of Puget Sound, and Harvard Vanguard Health System, may serve as a model for its capacity to implement studies and capture data within the context of usual care.⁵⁹ Other examples include AHRQ's 55 practice-based research networks, the VA Cooperative Studies Program, and a number of specialty-specific research networks sponsored by various NIH institutes. The NIH also supports the General Clinical Research Centers, a robust infrastructure for clinical research within academic medical centers. Additional attention to expanding this infrastructure beyond teaching hospital and clinics and to extending the capacity for clinical research into real-world clinical care settings would be extremely helpful in facilitating the conduct of clinical trials of value to decision makers.

It will also be necessary to significantly expand the number of physicians capable of conducting PCTs. The NIH and AHRQ should expand their programs to provide special career awards, scholarship, and loan-repayment programs aimed at expanding the clinical research workforce. Although some individuals will need extensive training to become principal investigators and devote their professional careers to these studies, a minimum level of skills in clinical investigation should become part of the basic training of physicians. Ultimately, with the expansion of electronic medical record systems, most clinical encounters could provide useful data and most physicians could function in part as clinical investigators.

Address Methodological and Ethical Issues

Practical clinical trials pose a number of methodological and ethical challenges

that will need to be identified and addressed. Clinical trialists have developed some strategies to reduce the cost of conducting large lengthy trials. Large simple trials feature simple protocols, enrollment from a large number of research sites, limited patient exclusion criteria, and data collection limited to the smallest possible number of elements.⁶⁰ This approach can make it possible to study thousands of patients at relatively low cost and was first shown to be feasible by Richard Peto and his colleagues, who conducted several large landmark trials of drug treatment for acute myocardial infarction.⁶¹ Broader application of these strategies to PCTs will be needed to minimize cost and participant burden. Because ancillary patient management in a PCT may not be standardized, there are likely to be clustering effects at the level of physicians and organizations that will require specialized methods to facilitate analysis or randomization at the organizational level. Such study designs pose some unusual methodological challenges that will require additional attention to refine.

For questions of comparative effectiveness, the expected outcome differences between 2 active therapies are likely to be considerably smaller than the differences between a treatment and placebo. Trials designed to demonstrate such differences will require large sample sizes and may be expensive. In some circumstances, fewer patients may be necessary to compare 2 active treatments in an equivalence trial: a study designed to demonstrate that there are no important differences in effectiveness between 2 active treatments. However, equivalence trials raise a number of challenging scientific and ethical issues and are not yet widely accepted, so additional conceptual development of this approach will be necessary.⁶²

Several important ethical issues may be encountered in the design and conduct of some PCTs. Because a number of these trials may involve comparison of new vs old technology, expensive vs inexpensive drugs, or high-intensity vs low-intensity services, there may be concerns about the economic motivations for

the trials and whether some patients enrolled in the trials are being denied optimal care (ie, whether true clinical equipoise exists).⁶³ Informed consent and confidentiality issues may be raised by more extensive use of clinical and administrative records to identify potential study participants and to augment the data collected specifically for research purposes. Consent procedures that are both adequate and efficient will need further exploration, because the number of PCTs that can be performed may be limited by the operational burden they impose, particularly in settings not primarily focused on research. In addition, efficient strategies for obtaining institutional review board approval from multiple institutions will be necessary, perhaps through approaches that involve coordination of central and local institutional review board reviews.

Funding Options for PCTs

A significant amount of funding will be necessary to support PCTs, given the large number and high cost of such trials. Furthermore, a substantial amount of support will be needed to develop the infrastructure necessary to efficiently conduct these studies. Resources to fund these trials will necessarily come either from redistribution of current research funds to important PCTs or by finding additional sources of new funds. Several possible public and private sources of additional or redistributed funding should be considered.

Although increasing the portion of NIH resources flowing to PCTs may have merit, given the NIH mission to improve public health through research, there is likely to be concern that the money is being diverted from important basic research. However, the NIH has recently acknowledged the need for improvements in the clinical research enterprise⁶⁴ and may be increasingly willing to collaborate with decision makers once they have more clearly identified their critical clinical research priorities. Funds from the NIH are also needed to support further development of the PCT infrastructure, including support for research networks, expansion of the Gen-

eral Clinical Research Centers beyond academic centers, and training physician-investigators who have skills to design and conduct PCTs.

Industry funding could also support PCTs. This is most likely to occur in response to the increased use of evidence-based policy development by decision makers. Industry will be particularly responsive to coverage and payment policy firmly linked to evidence of improved health outcomes and increased payment to evidence of substantially improved outcomes. The regulatory framework of the FDA is designed to ensure that medical technologies approved for marketing are safe and effective for their intended use and is not structured to ensure that research is conducted that will inform optimal clinical use of these technologies. Additional requirements for comparative studies before FDA approval could significantly increase the time and expense of bringing new technologies to market. Recent policy discussions have been focused on increasing the speed of FDA approvals by developing strategies for gathering additional data in the postmarketing period.⁶⁵ Increased diffusion of health information technology will facilitate more rapid and efficient conduct of requisite comparative studies, in both the premarket and postmarket phases.

Because payers and purchasers increasingly demand high-quality evidence to support decisions, they are often proposed as the most likely source of additional funding for PCTs. The benefit of an enhanced evidence base for making clinical and health policy decisions ultimately accrues to those who use health care services. For that reason, financial support and priority setting for this body of research would most sensibly be associated with the funding streams that support the provision of health care services. Furthermore, the amount of funding for this body of applied research should be linked to the level of resources being consumed in supplying health care services.

Several health care delivery systems allocate dedicated resources for clinical research. The VA spends approxi-

mately \$24 billion on the delivery of health care services and maintains a budget of \$350 million for research, including approximately \$60 million for clinical trials (virtually all of which are PCTs). The National Health Services in England and Canada also have provided funding for a large number of PCTs. These are examples of health systems where health care services are delivered within a fixed budget and clinical research is understood to be a critical mechanism to ensure optimal use of existing resources. Such an approach should be considered in the broader US health care systems and could provide substantial resources for high-priority applied clinical research if even a small fraction of total health care spending were reserved for this purpose.

This proposal, although conceptually simple, may be challenging to implement. Most private insurance contracts exclude payment for experimental and investigational services, although coverage of clinical trials is increasingly popular. At present, the Medicare program does not have broad legal authority to fund clinical research, because the statute does not allow payment for care unless medically necessary. If payers and purchasers were to identify funds for clinical research, it would be essential to have a functioning system for priority setting free of political interference and robust mechanisms for oversight of sponsored research. Finally, costs for supporting clinical research will, at least initially, increase overall health care spending or replace spending now assigned to other technologies and services. Although PCTs might be expected to increase the efficiency of spending and the health benefit derived from given dollars, any increased costs for clinical research will eventually be passed on to the payers, purchaser, employers, public programs, and eventually the general public. Health care consumers will ultimately need to decide whether this is a desirable use of their money.

CONCLUSION

The United States is now seeing the result of its heavy investment in biomed-

ical research: numerous promising medical products have been developed and many more are on the way to initial clinical trials. With this success comes an equally important additional need: to develop a systematic approach to efficiently evaluate the risks and benefits of these new technologies in the context of existing alternatives. Currently, decision makers in health care are not provided with information of adequate quality to make well-informed decisions regarding alternative strategies for diagnosis and treatment of most common clinical conditions. Improving the quality of clinical research will depend on more active involvement of clinical and health policy decision makers in all aspects of clinical research, including priority setting, study design, study implementation, and funding.⁶⁶

Disclaimer: The opinions contained in this article represent the views of the individual authors and do not necessarily reflect the views or policies of the Centers for Medicare & Medicaid Services or AHRQ.

Acknowledgment: We thank Joe Selby, MD, Allan Korn, MD, and Robert Califf, MD, for comments on earlier drafts of the manuscript. Alex Omay, PhD, and the members of the Institute of Medicine Clinical Research Roundtable contributed greatly through many insightful discussions of these issues. We also acknowledge the wisdom and support of Peter Mazonson, MD, in the earliest stages of this work.

REFERENCES

1. Wennberg J, Gittelsohn A. Small area variation in health care delivery. *Sci Am*. 1973;182:1102-1108.
2. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? *Milbank Q*. 1998;76:517-563.
3. Eddy DM, Billings J. The quality of medical evidence: implications for quality of care. *Health Aff (Millwood)*. 1988;7:19-32.
4. McNeil BJ. Shattuck lecture—hidden barriers to improvement in the quality of care. *N Engl J Med*. 2001;345:1612-1620.
5. Field MJ, Lohr KN, eds. *Clinical Practice Guidelines: Directions for a New Program*. Washington, DC: National Academy Press; 1990.
6. Steiner CA, Powe NR, Anderson GF, Das A. Technology coverage decisions by health care plans and considerations by medical directors. *Med Care*. 1997;35:472-489.
7. Garber AM. Evidence-based coverage policy. *Health Aff (Millwood)*. 2001;20:62-82.
8. Antman K, Lagakos S, Drazen J. Designing and funding clinical trials of novel therapies. *N Engl J Med*. 2001;344:762-763.
9. Altman SH, Parks-Thomas C. Controlling spending for prescription drugs. *N Engl J Med*. 2002;346:855-856.
10. Sung NS, Crowley WF Jr, Genel M, et al. Central challenges facing the national clinical research enterprise. *JAMA*. 2003;289:1278-1287.
11. Mulrow CD, Williams JW Jr, Trivedi M, et al. *Treatment of Depression: Newer Pharmacotherapies*. Rock-

- ville, Md: Agency for Health Care Policy and Research; 1999. AHCPR publication 99-E014.
12. Lomas J, Anderson G, Enkin M, et al. The role of evidence in the consensus process: results from a Canadian consensus exercise. *JAMA*. 1988;259:3001-3005.
 13. Bach PB, Brown C, Gelfand SE, McCrory DC. Management of acute exacerbations of chronic obstructive pulmonary disease: a summary and appraisal of published evidence. *Ann Intern Med*. 2001;134:600-620.
 14. Snow V, Lascher S, Mottur-Pilson C, for the Joint Expert Panel on Chronic Obstructive Pulmonary Disease of the American College of Chest Physicians and the American College of Physicians-American Society of Internal Medicine. Evidence base for management of acute exacerbations of chronic obstructive pulmonary disease. *Ann Intern Med*. 2001;134:595-599.
 15. Tunis SR, Kang JL. Improvements in Medicare coverage of new technology. *Health Aff (Millwood)*. 2001;20:83-85.
 16. National Institutes of Health, Grants-OER Home Page. Minimally invasive surgical therapies treatment consortium for benign prostatic hyperplasia. Available at: <http://grants.nih.gov/grants/guide/ffa-files/RFA-DK-01-024.html>. Accessed July 5, 2002.
 17. Centers for Medicare and Medicaid Coverage. Air-fluidized beds for pressure ulcers. Available at: <http://www.cms.hhs.gov/coverage/8b3-q6.asp>. Accessed July 5, 2002.
 18. Rafuse J. Education, practice reviews needed to reduce surgical intervention. *CMAJ*. 1996;155:463-464.
 19. Connors AF, Speroff T, Dawson NV, et al, for the SUPPORT Investigators. The effectiveness of right heart catheterization in the initial care of critically ill patients. *JAMA*. 1996;276:889-897.
 20. Soumerai SB, McLaughlin TJ, Spiegelman D, et al. Adverse outcomes of underuse of β -blockers in elderly survivors of acute myocardial infarction. *JAMA*. 1997;277:115-121.
 21. Naylor CD, Guyatt GH, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, X: how to use an article reporting variation in the outcomes of health services. *JAMA*. 1996;275:554-558.
 22. Greenland S. Randomization, statistics, and causal inference. *Epidemiology*. 1990;1:421-429.
 23. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288:321-333.
 24. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis*. 1967;20:637-648.
 25. Roland M, Torgerson DJ. What are pragmatic trials? *BMJ*. 1998;316:285.
 26. Lepor H, Williford WO, Barry MJ, et al. The efficacy of terazosin, finasteride, or both in benign prostatic hyperplasia. *N Engl J Med*. 1996;335:533-540.
 27. The MTOPS Research Group. The impact of medical therapy on the clinical progression of BPH: results of the MTOPS trial. Paper presented at: American Urological Association Annual Meeting; May 28, 2002; Orlando, Fla. Available at: <http://www.niddk.nih.gov/welcome/releases/05-28-02abstract.pdf>. Accessed July 25, 2002.
 28. Cherkin DC, Deyo RA, Battie M, et al. A comparison of physical therapy, chiropractic manipulation, and provision of an educational booklet for the treatment of patients with low back pain. *N Engl J Med*. 1998;339:1021-1029.
 29. Cherkin DC, Eisenberg D, Sherman KJ, et al. Randomized trial comparing traditional Chinese medical acupuncture, therapeutic massage, and self-care education for chronic low back pain. *Arch Intern Med*. 2001;161:1081-1088.
 30. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, II: how to use an article about therapy or prevention, B: what were the results and how will they help me in caring for my patients? *JAMA*. 1994;271:59-63.
 31. Dans AL, Dans LF, Guyatt GH, Richardson S, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, XIV: how to decide on the applicability of clinical trials results to your patient. *JAMA*. 1998;279:545-549.
 32. Califf RM, DeMets DL. Principles from clinical trials relevant to clinical practice: part I. *Circulation*. 2002;106:1015-1021.
 33. Williams JW, Barrett J, Oxman T, et al. Treatment of dysthymia and mild depression in primary care: a randomized controlled trial in older adults. *JAMA*. 2000;284:1519-1526.
 34. Dolor RJ, Witsell DL, Hellkamp AS, et al. Comparison of cefuroxime with or without intranasal fluticasone for treatment of rhinosinusitis: the CAFFSTrial: a randomized controlled trial. *JAMA*. 2001;286:3097-3105.
 35. Delaney BC, Wilson S, Roalfe A, et al. Randomised controlled trial of *Helicobacter pylori* testing and endoscopy for dyspepsia in primary care. *BMJ*. 2001;322:898-901.
 36. Grimm RH Jr, Margolis KL, Papamertziou VV, et al. Baseline characteristics of participants in the Antihypertensive and Lipid Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *Hypertension*. 2001;37:19-27.
 37. The ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in high-risk patients randomized to angiotensin-converting enzyme inhibitor or calcium-channel blocker vs diuretic: the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *JAMA*. 2002;288:2981-2997.
 38. ALLHAT Collaborative Research Group. Major cardiovascular events in hypertensive patients randomized to doxazosin vs chlorthalidone: the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *JAMA*. 2000;283:1967-1975.
 39. Guyatt GH, Rennie D. The basics: using the medical literature. In: *Users' Guides to the Medical Literature*. Chicago, Ill: AMA Press; 2002.
 40. Barry MJ, Cockett AT, Holtgrewe HL, et al. Relationship of symptoms of prostatism to commonly used physiological and anatomical measures of the severity of benign prostatic hyperplasia. *J Urol*. 1993;150:351-358.
 41. Lederle FA, Wilson SE, Johnson GR, Reinke DB. Immediate repair compared with surveillance of small abdominal aortic aneurysms. *N Engl J Med*. 2002;346:1437-1444.
 42. The United Kingdom Aneurysm Trial Participants. Long-term outcomes of immediate repair compared with surveillance of small abdominal aortic aneurysms. *N Engl J Med*. 2002;346:1445-1452.
 43. Rizzo JD, Powe NR. Methodological hurdles in conducting pharmacoeconomic studies. *Pharmacoeconomics*. 1999;15:339-355.
 44. Simon GE, VonKorff M, Heiligenstein JH, et al. Initial antidepressant choice in primary care: effectiveness and cost of fluoxetine vs tricyclic antidepressants. *JAMA*. 1996;275:1897-1902.
 45. Snow V, Lascher S, Mottur-Pilson C, for the American College of Physicians-American Society of Internal Medicine. Pharmacologic treatment of acute major depression and dysthymia: clinical guideline, part 1. *Ann Intern Med*. 2000;132:738-742.
 46. National Institutes of Health. St John's wort study in mild depression to enroll 300 participants. *Health News Daily*. 2003;15:5-6.
 47. Nathan DG, Varmus HE. The National Institutes of Health and clinical research: a progress report. *Nat Med*. 2000;6:1201-1204.
 48. Schein OD, Katz J, Bass EB, et al. The value of routine preoperative medical testing before cataract surgery: study of medical testing for cataract surgery. *N Engl J Med*. 2000;342:168-175.
 49. Fitzgibbons RJ. *Grants On-Line Database* [database online]. Abstract of grant HS09860: inguinal hernia management: watchful waiting vs operation. Agency for Healthcare Research and Quality. Available at: <http://www.gold.ahrq.gov/index.cfm>. Accessed July 20, 2002.
 50. Woosley RL. Centers for Education and Research in Therapeutics. *Clin Pharmacol Ther*. 1994;55:249-255.
 51. Takaro T, Hultgren HN, Lipton MJ, Detre KM. The VA cooperative randomized study of surgery for coronary arterial occlusive disease II: subgroup with significant left main lesions. *Circulation*. 1976;54(suppl 6):107-117.
 52. Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med*. 2002;347:81-88.
 53. Kolata G. VA suggests halt to kind of knee surgery. *New York Times*. August 24, 2002;A:9(column 4).
 54. Getz K, Zisson S. Clinical grants market decelerates. *CenterWatch*. 2003;10:4-5.
 55. Sandham JD, Hull RD, Brant RF, et al. A randomized, controlled trial of the use of pulmonary-artery catheters in high-risk surgical patients. *N Engl J Med*. 2003;348:5-14.
 56. Siris ES, Miller PD, Barrett-Connor E, Faulkner KG. Identification and fracture outcomes of undiagnosed low bone mineral density in postmenopausal women: results from the National Osteoporosis Risk Assessment. *JAMA*. 2001;286:2815-2822.
 57. Tunis SR, Korn A, Omayra A. *The Role of Purchasers and Payers in the Clinical Research Enterprise*. Washington, DC: National Academy Press; 2002.
 58. Mello MM, Brennan TA. The controversy over high-dose chemotherapy with autologous bone marrow transplant for breast cancer. *Health Aff (Millwood)*. 2001;20:101-117.
 59. HMO Research Network. HMO Research Network Home Page. Available at: <http://hmoresearch-network.org/>. Accessed June 21, 2002.
 60. Buring JE, Jonas MA, Hennekens CH. Large and simple randomized trials. In: US Congress, Office of Technology Assessment. *Tools for Evaluating Health Technologies: Five Background Papers*. Washington, DC: Office of Technology Assessment; 1995.
 61. Baigent C, Collins R, Appleby P, Parish S, Sleight P, Peto R, for the ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. ISIS-2: 10 year survival among patients with suspected acute myocardial infarction in randomised comparison of intravenous streptokinase, oral aspirin, both, or neither. *BMJ*. 1998;316:1337-1343.
 62. Djulbegovic B, Clarke M. Scientific and ethical issues in equivalence trials. *JAMA*. 2001;285:1206-1208.
 63. Liliford R. Ethics of clinical trials from a bayesian and decision analytic perspective: whose equipoise is it anyway? *BMJ*. 2003;326:980-981.
 64. DeAngelis CD. NIH Director Elias A. Zerhouni, MD, reflects on agency's challenges, priorities. *JAMA*. 2003;289:1492-1493.
 65. US Food and Drug Administration. Improving innovation in medical technology: beyond 2002. Available at: <http://www.fda.gov/bbs/topics/NEWS/2003/beyond2002/report.html>. Accessed April 30, 2003.
 66. Califf RM, DeMets DL. Principles from clinical trials relevant to clinical practice: part II. *Circulation*. 2002;106:1172-1175.