ABSTRACT The Patient Protection and Affordable Care Act established a new Patient-Centered Outcomes Research Institute to identify and address research priorities for comparative effectiveness research. Among its many responsibilities, the institute has been charged with setting priorities, developing methodological standards, and communicating research results to decision makers. In this paper we consider how the institute can support the different standards for acceptable evidence used by various government agencies, providers, patients, and other decision makers. We argue that the public interest will best be served if the institute develops a balanced and flexible approach to deciding which types of studies to fund.

The Patient Protection and Affordable Care Act of 2010 established the Patient-Centered Outcomes Research Institute to identify and address research priorities for comparative effectiveness research.1 That research will be used by a range of decision makers, including government regulators determining whether to license a new medical product, policy makers and payers deciding what interventions health plans should cover, and individual providers and patients choosing treatments. Should the new institute strive to have all of the research it sponsors meet a specific standard for acceptable evidence—a so-called evidentiary standard—even though the research will be used by different decision makers, whose standards vary?

The health reform law does not mention evidentiary standards explicitly. However, it does mandate developing and periodically updating scientifically based methodological standards—or the standards for conducting different types of research, such as clinical trials. The law also requires the appropriate use of comparative effectiveness research in Medicare’s coverage, reimbursement, and incentive programs.

The use of comparative effectiveness research results in policy decisions raises the question of whether evidentiary standards for the new research must be consistent with other governmental standards. The Affordable Care Act recognizes that comparative effectiveness research will include systematic reviews and observational studies, as well as randomized controlled trials.

Evidence from randomized controlled trials is generally considered to be the most reliable. As a result, the trials are the gold standard for medical evidence, preferred by the Food and Drug Administration (FDA) in deciding whether to approve new drugs, for instance.

In this paper we discuss the existing evidentiary standards of government agencies and their implications for the standards for comparative effectiveness research conducted under the auspices of the new health care law. The Patient-Centered Outcomes Research Institute is not charged with establishing the evidentiary standards that decision makers—be they patients, providers, regulators, or payers—will apply to the results generated by research. We argue that the institute should, as mandated, focus on methodological standards and the best scientific...
practice for each type of comparative effectiveness research, rather than on evidentiary standards for use of the research.

### Decision Makers’ Needs

Comparative effectiveness research, as defined for the institute by the Affordable Care Act, means “research evaluating and comparing health outcomes and the clinical effectiveness, risks, and benefits of 2 or more medical treatments, services, and items.” The research can assess health care interventions such as procedures, diagnostic tests, pharmaceuticals, medical devices, and other technologies.

Three key points define how it differs from other types of health care research. Comparative effectiveness research compares two or more alternative interventions; focuses on effectiveness, or real-world outcomes, as opposed to efficacy, or experimental outcomes; and aims to provide information to a wide range of decision makers, including patients, providers, and policy makers.

In Appendix Exhibit 1, we trace the relationships among various decision makers, the issues they need to decide, and the research’s data and evidence, as well as the dissemination of the results. Most of the relationships are obvious and self-explanatory, but it is worth emphasizing a few key points.

First, many decision makers need comparative effectiveness results to inform their decisions. Each group is likely to rely on different methods and evidence in making decisions.

Second, within comparative effectiveness research, data become evidence only after analytic processing and interpretation. For example, data from electronic medical records and related medical claims must be linked, organized into treatment subgroups, and then compared across subgroups in order to generate evidence on how different treatments affect the cost of care.

Third, uncertainty is pervasive and persistent, and one area of uncertainty is how much evidence is needed. Marketing approval by regulators, product selection by physicians and patients, and coverage by payers, for instance, all require different amounts and types of evidence. This variation gives rise to the perpetual question about the benefits versus the costs of more research, or what the additional research’s “value” is.

Fourth, comparative effectiveness research can assess many different types of medical technologies, including drugs, procedures, and even health systems. Different types of decision makers face different policy issues, and the evidentiary standards they apply vary not only among types of decision makers but also across types of medical technologies.

For example, the FDA might need to decide whether to approve a new technology—a drug, device, or procedure—for sale, while a provider might have to decide whether to recommend a certain technology for patient use. Their evidentiary standards might be quite dissimilar.

### FDA Evidentiary Standards

Across the federal government, different standards for the assessment of evidence are set out in statutes, regulations, and policies. Even within a single department or agency, evidentiary standards may vary depending on what is being assessed and the purpose of the assessment.

For example, the FDA has multiple standards for reviewing evidence (Appendix Exhibit 2). The standard for approval of new drugs and biologicals is substantial evidence. In general, at least two major trials are required for product approval, although since 1997, one trial with supporting evidence is sufficient in limited circumstances. In contrast, the standard for approval of a Class III, or high-risk, medical device is reasonable assurance that the device is safe and effective.

There is, thus, some flexibility in how the standards are applied to specific products and types of products. The variations are an acknowledgment of the need to ensure public health and safety while also considering the feasibility, timeliness, and cost of obtaining evidence about a product before it is introduced commercially.

For example, the FDA’s evidentiary standard for allowing a new drug onto the market is usually two double-blind, randomized trials. In contrast, following the end of a drug’s patent protection or other market exclusivity, a generic copy needs only to demonstrate equivalent chemical and biological activity to the first drug. Similarly, the standard for a new device, such as an artificial hip, depends on its similarity to an existing approved device. Additional safety data are required only if the new device is substantially different from previously approved devices.

Dissemination of information about an approved product is also highly regulated by the FDA, particularly in connection to statements made or implied by the product’s manufacturer. Prescription drug advertising for approved uses—within-label uses rather than off-label uses—must meet the standard of substantial clinical experience or substantial evidence, which requires controlled clinical trials to support claims about the drug.
However, evidentiary standards are lower for data that suggest the presence of previously unrecognized safety issues. Annual safety reports after a product has entered the marketplace often include only epidemiological or observational studies, rather than a more rigorous clinical trial.

A different standard—competent and reliable scientific evidence—applies to a drug company’s claim to a health plan about a product’s value and effectiveness. This standard was adopted in 1997 to meet the need of managed care plans and other payers for information such as whether a diabetes medication saves money overall because it prevents disease-related complications.

As discussed below, the Federal Trade Commission (FTC) uses the same language in its evidentiary standards, but it is unclear whether the two agencies interpret the standard in the same way.

In practice, the FDA’s application of evidentiary standards depends on judgments related to the novelty of the product, the medical need for new therapies for the target condition, and what is known about the product’s effectiveness relative to its risks. The agency is generally more willing to approve a product with fewer data when there is a substantial medical need for new therapies that could alleviate serious harm posed by the target disorder. More evidence is usually required when a new product seems to offer little advantage over established treatments or when there are unanswered questions about the product’s safety.

Standards Of Other Agencies
Evidentiary standards used by other government agencies involve different legal frameworks (Appendix Exhibit 3). For example, for coverage of items and services under Medicare, the Centers for Medicare and Medicaid Services (CMS) has set a standard of reasonable and necessary. Although CMS has not succeeded in issuing regulations that define this standard, it has issued guidance to help explain its intent regarding the standard.

CMS has also issued decision memoranda explaining the evidence it considers when it decides whether to cover items and services. It is clear that the reasonable and necessary standard is intended to be different from the FDA’s standard for approving a product’s marketing. CMS uses the standard to determine coverage for FDA-approved products, including off-label uses. For example, drugs approved by the FDA for use in lymphoma are covered by CMS for use in colorectal cancer (an off-label use). At the same time, CMS has refused to cover some medical devices and related procedures—such as computed tomography (CT) colonography—despite their FDA approval.

The Federal Trade Commission shares jurisdiction with the FDA over advertising claims and has sole jurisdiction over promotion of medical procedures that do not involve devices regulated by the FDA. The commission employs the standard of competent and reliable scientific evidence in assessing whether advertising claims are warranted. This standard requires evidence consistent with what experts in the relevant field would consider appropriate to support claims. Thus, this truth-in-advertising standard is flexible and can reflect changes in the scientific state of the art.

Key Issues And Implications
Evidence generated by comparative effectiveness research will provide benefits for many different decision makers. In that sense, the research is a public good. Indeed, it is useful to think about the research in terms of what economists call the “value of information,” which is the aggregate of what each of the decision makers would be willing to pay for it, as compared to its cost.

The overall benefit derived from dissemination of comparative effectiveness research will be a better understanding of the impact of treatment decisions on outcomes, such as improvements in patients’ life expectancy and illness levels. In setting the agenda for research priorities, the Patient-Centered Outcomes Research Institute should consider the research’s relative value, defined as “based on the cost of conducting research compared to the potential usefulness of the information.”

Economists can construct mathematical models to estimate the financial value of gathering more information. These complex models typically rely on a wide range of evidence about key factors, including the health benefits (efficacy) of new treatments based on randomized clinical trials and the relationship between that short-term efficacy and outcomes that might occur many years later, such as a heart attack, based on epidemiological studies. In setting its research project agenda, the institute is directed to consider value-of-information analysis, which uses models to synthesize a wide range of evidence to inform decision making.

THE ROLE OF MODELS Different decision makers draw on various types of evidence and sources, and organize that information in some sort of framework—an implicit model. For example, FDA drug advisory committees are asked to judge whether the expected benefits of a new drug outweigh its expected risks. These models
that support decision making often apply different standards in different situations, and most are eclectic in the way they use available information. These models are flexible and updated continually as new information becomes available.

Sir Michael Rawlins, the head of the National Institute for Health and Clinical Excellence (NICE) in England and Wales, has commented: “Experiment, observation, and mathematics, individually and collectively, have a crucial role in providing the evidential basis for modern therapeutics. Arguments about the relative importance of each are an unnecessary distraction. Hierarchies of evidence should be replaced by accepting—indeed embracing—a diversity of approaches.” Indeed, some experts have identified decision modeling—the use of complex mathematical models to integrate different types of evidence—as a method of comparative effectiveness research.5

As noted above, two well-controlled clinical trials demonstrating efficacy are often what the FDA requires to authorize the marketing of a drug. In practice, however, the agency’s decisions are based on the trials’ evidence and the interpretation and judgment of experts, both external and internal. This reality is reflected in the often-heard quip that “evidence doesn’t make decisions, people do.”

THE EVIDENCE NEEDS OF REGULATORS AND PAYERS The emergence of government-sponsored comparative effectiveness research raises questions about evidentiary standards across and within government agencies.6 Standards differ between reimbursement and regulatory agencies and, as noted above, for different products and purposes.7 This is not surprising, given that different agencies have different objectives.

Although variety in evidentiary standards has always existed, the fact that the federal government is holding groups accountable to different yardsticks is highlighted by the increased resources flowing to comparative effectiveness research. This inconsistency sends different messages to patients and providers, who ultimately must make treatment decisions based on different types of evidence.

The evidence itself is subject to legal limitations on the dissemination of information. For example, manufacturers of products regulated by the FDA are prohibited from disseminating promotionally or other proactive communications about a product that is inconsistent with the package labeling. But the manufacturers are permitted to communicate such information in response to a specific request for it. And a payer may disseminate information about the off-label use of the same product.

Some analysts have argued that following principles of comparative effectiveness, the FDA should require studies comparing a new drug to the next-best alternative instead of traditional controlled trials using placebos.4 These analysts contend that current regulations result in the use of drugs with no supporting evidence of superiority over alternatives and that physicians and other decision makers need more clinically relevant information when new drugs reach the market. But this desire for more data must be balanced with the public health implications of delayed approval of new treatments.

Other experts have recommended that manufacturers be required to provide comparative effectiveness data prior to the widespread adoption of a new treatment, or required to report on the label and in marketing materials what is known about the comparative effectiveness of a treatment.3 These recommendations envision that the FDA will be the arbiter of evidence to be used to inform clinical decision making.

Such new regulations could have unintended effects. The FDA plays an economic role as a market regulator in addition to its prescribed role as protector of the public health. If multiple products provide adequate evidence of safety and effectiveness for a particular condition, the FDA is expected to license all of the products. The presence of multiple products increases competition in the marketplace and tends to drive down prices, which can benefit the public.

Moreover, the benefits of products are often enhanced over time as new information becomes available based on biological, genomic, and environmental factors demonstrating that one treatment may be more effective or safer than others for certain subgroups or even individual patients. The presence of new comparative effectiveness evidence requirements that might support limiting the number of products available to the marketplace at the outset would reduce these potential benefits.

A BALANCED AGENDA FOR RESEARCH Because comparative effectiveness evidence will increasingly be generated and disseminated through publicly sponsored research in a new, nonregulatory environment, the appropriate communication and interpretation of the evidence will be critical. The Patient-Centered Outcomes Research Institute intends to release research findings and information about their limitations—including what further research may be needed—in a form that clinicians and the general public can use in making health care decisions.

This information may not include practice guidelines or recommendations about coverage or payment. Yet it could be used by various de-
cision makers in formulating practice guidelines and determining what interventions to cover and how much to pay for them.

The institute is not required to communicate information in a manner consistent with FDA-approved labeling of regulated products. Federal payers are not prohibited from using research findings to inform coverage, payment, and treatment decisions if the findings are used in conjunction with other evidence.

As a result, the research sponsored by the institute may communicate information that includes a more direct assessment of value than some decision makers, such as product manufacturers, may be authorized to publicly address. For example, under current rules, a manufacturer may not claim that its product is better than a competing treatment, although the manufacturer and the institute, whose research may make that comparative effectiveness finding, rely on the same type of data to make their determinations.

It will be essential for the institute to develop a balanced research agenda and to identify the appropriate context for effectively communicating research findings. Consistent frameworks and methods of communication, at least to the public, will be needed if the patient—the chief intended beneficiary of the information—is to be able to make practical use of it.

Conclusion
The Affordable Care Act wisely calls for the Patient-Centered Outcomes Research Institute’s methodological standards to “provide specific criteria for internal validity, generalizability, feasibility, and timeliness of research,” recognizing the need for balance among these objectives. The law also directs the institute to develop and update “a translation table...designed to provide guidance and act as a reference for the Board to determine research methods that are most likely to address each specific research question.”

It is important to note that the law is addressing methodological standards, not evidentiary standards. Methodological standards have to do with the best practice for a given type of methodology, be it a clinical trial, an observational study, or the analysis of data from either of these types of studies.

Evidentiary standards pertain to decision makers’ choices, such as whether to require data from controlled trials. The Affordable Care Act does not call for rigid evidentiary standards; nor does it assign the institute the task of developing and updating such standards.

To maximize the value of the information it produces, the Patient-Centered Outcomes Research Institute should take a balanced and flexible approach to the types of research it sponsors, being careful not to let rigorous scientific methods become a rigid evidentiary standard.

This research was supported by an unrestricted grant from the National Pharmaceutical Council to the University of Washington. The authors thank their Pharmaceutical Council to the University unrestricted grant from the National sponsors, being careful not to let rigorous scientif

experts who discussed the key issues and shared their insights in background interviews: Marc Berger, Patrick Conway, Cliff Goodman, Scott Gottlieb, Bryan Luce, Steve Pearson, Steve Phurrough, Jean Slutsky, Sean Tunis, and Gail Wilensky. Finally, the authors thank Jennifer Graf and Les Paul of the National Pharmaceutical Council for their support of and comments on a draft report and the manuscript. All opinions expressed are those of the authors alone, and any errors are the authors’ responsibility.

NOTES
1 The Patient Protection and Affordable Care Act of 2010. PL 111-148, sec. 6301.
2 To access the Appendix, click on the Appendix link in the box to the right of the article online.
9 O’Connor A. Building comparative efficacy and tolerability into the FDA approval process. JAMA. 2010;303:979–80.
Different users of comparative effectiveness research will employ different standards of evidence to make decisions—and that is how it should be, write Lou Garrison and colleagues. Thus, the new Patient-Centered Outcomes Research Institute doesn’t need to concern itself with refining these evidentiary standards. It can focus instead on other critical areas, such as methods for conducting the research.

Garrison is an economist and professor in the Department of Pharmacy at the University of Washington, in Seattle, where he teaches pharmacy policy methods. His research focuses on economic evaluation of drugs and diagnostics used in organ transplantation, renal disease, flu, and cancer.

Before turning to teaching, Garrison worked for twelve years as an economist in the pharmaceutical industry, as vice president and head of health economics and strategic pricing at Roche Pharmaceuticals based in Basel, Switzerland. Earlier he was director of Project HOPE’s Center for Health Affairs, where he researched the impact of technological change on health care costs. Garrison earned a doctorate in economics from Stanford University.

Peter Neumann is a professor of medicine and director of the Center for the Evaluation of Value and Risk in Health, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, and a professor of medicine at Tufts University School of Medicine. He was on the faculty of the Harvard School of Public Health for ten years, most recently as associate professor of policy and decision sciences.

Neumann’s research focuses on the use of cost-effectiveness analysis in health care decision making, including ways to value willingness to pay for care. He also founded and directs the Cost-Effectiveness Registry (www.cearegistry.org), a database of cost-effectiveness analyses in health care, and is a member of the Health Affairs editorial board.

Neumann has served in policy positions in Washington, D.C., as president of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), and as a trustee of the Society for Medical Decision Making. He received a doctorate in health policy and management from Harvard University.

Paul Radensky is partner in the law firm of McDermott Will and Emery LLP, in Washington, D.C., and Miami, and a board-certified internal medicine physician. He advises clients on pharmaceutical, biotechnology, medical device, and clinical laboratory development and marketing from product design and development through Food and Drug Administration (FDA) approval, manufacturing, and payer reimbursement. He earned his medical degree from the University of Pennsylvania and his law degree from Harvard Law School.

Sheila Walcoff is a partner in the law firm of McDermott Will and Emery LLP, in Washington, D.C., where her practice focuses on personalized medicine and federal regulatory and science policy issues. Previously, she served as counselor in the US Department of Health and Human Services, Office of the Secretary, and as associate commissioner for external relations at the FDA, where she provided policy and communications counsel. Walcoff earned her law degree from the Georgetown University School of Law.