This product is part of the RAND Corporation technical report series. Reports may include research findings on a specific topic that is limited in scope; present discussions of the methodology employed in research; provide literature reviews, survey instruments, modeling exercises, guidelines for practitioners and research professionals, and supporting documentation; or deliver preliminary findings. All RAND reports undergo rigorous peer review to ensure that they meet high standards for research quality and objectivity.

TECHNICAL REPORT

Dissemination and Adoption of Comparative Effectiveness Research Findings When Findings Challenge Current Practices

Eric C. Schneider, Justin W. Timbie, D. Steven Fox, Kristin R. Van Busum, John P. Caloyeras

Sponsored by the U.S. Department of Health and Human Services

RAND HEALTH

The research described in this report was sponsored by the U.S. Department of Health and Human Services and was conducted in RAND Health, a division of the RAND Corporation.

Published 2011 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: http://www.rand.org
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

Executive Summary

Background

Insufficient evidence regarding the effectiveness of medical treatments has been identified as a key source of inefficiency in the U.S. healthcare system. Clinicians vary widely in their recommendation and use of diagnostic tests and treatments for patients with similar symptoms or conditions. This variation has been attributed to clinical uncertainty, since the published scientific evidence base does not provide adequate information to determine which treatments are most effective for patients with specific clinical needs.

A dramatic federal investment in comparative effectiveness research (CER) was made possible through the American Recovery and Reinvestment Act of 2009 (ARRA), with the expectation that the results will not only influence clinical practice but will also improve the efficiency of healthcare delivery. To do this, CER must provide information that supports fundamental changes in healthcare delivery and informs the choice of diagnostic and treatment strategies. Many new tests and treatments commonly adopted today are not completely grounded in scientific evidence. Some remain entrenched even when unambiguous scientific evidence about superior alternative approaches emerges. Other new clinical practices are not quickly adopted, either because information about them does not reach decisionmakers in a usable format or because of other barriers to their adoption.

Study Objectives

The project described in this report had three main objectives: (1) to develop a framework to help organize the array of barriers and enablers that influence the translation of CER evidence into new clinical practices; (2) to conduct case studies on the adoption of new clinical practices; and (3) to identify policy options that might facilitate dissemination of CER-based clinical practices.

We designed our organizational framework to isolate key factors affecting each phase of the process of CER translation, beginning with the generation of evidence and ending with the adoption of new clinical practices. The framework was also intended to inform CER development and dissemination activities, as well as future research on translation of CER into practice.

We conducted case studies on the adoption of new clinical practices following the release of five carefully selected CER studies published in the past 15 years, applying our framework to identify key themes relating to the pace of adoption of new practices. We sought information through discussions with stakeholders representing a broad range of perspectives and by examining the peer-reviewed literature associated with each case study. Synthesizing common themes across case studies provided insight into the root causes for the failure of CER to change clinical practice in a timely manner.

We developed a set of policy options through consultation with an expert panel and with partners at the Office of the Assistant Secretary for Planning and Evaluation (ASPE) in the Department of Health and Human Services (HHS) that might facilitate dissemination of CER-based clinical practices and thus maximize the effectiveness of the federal government's current investment in CER.

The methodology we developed might also be used to inform larger-scale, prospective, in-depth qualitative and quantitative research on the impact of the federal investment in CER.

Methodology

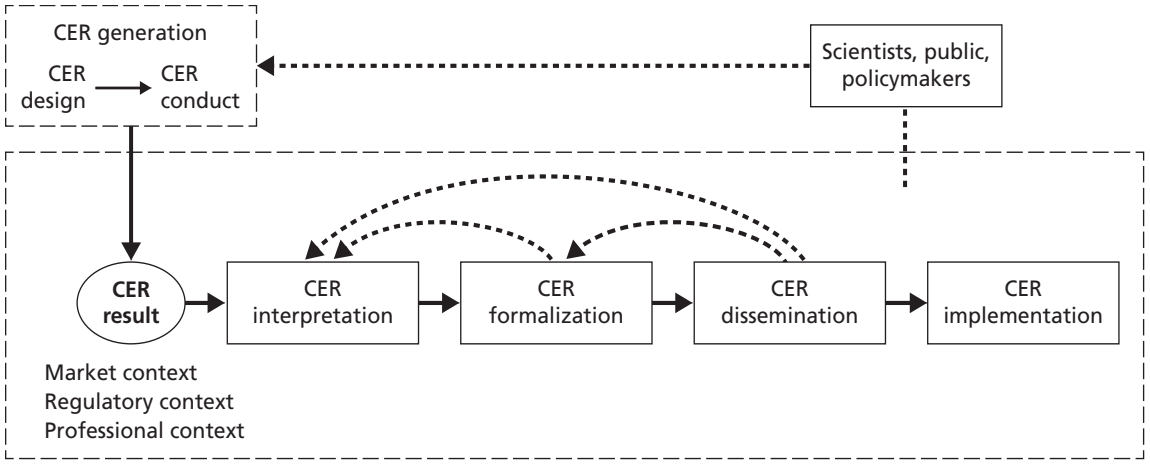
Framework for CER Translation into Practice

Our conceptual framework posits that the process of CER translation follows five key phases, shown in Figure S.1. While Figure S.1 suggests a generally linear temporal process, the phases are actually somewhat concurrent, and there appear to be multiple interactions between stakeholders at different phases. The phases are described in Table S.1.

Case-Study Research Approach

In selecting case studies, our preliminary intent was to identify and include CER trials that produced results that challenged current clinical practices. We used an environmental scan to develop a preliminary list of case-study topics, which we narrowed to five, based on a number of considerations. We wanted the topics to involve a high burden of illness, high prevalence, high-quality studies, diversity of treatment modalities, and diversity of treatment settings. The case-study topics chosen are shown in Table S.2.

Figure S.1
Conceptual Framework for Translation of CER into Clinical Practice



NOTES: CER dissemination and implementation phases involve dissemination and implementation of *clinical practices* as opposed to *results*. Solid lines represent the pathways through which CER research is translated into clinical practice; dotted lines are feedback loops that influence prior phases, indicating the interactive and iterative nature of the process.

RAND *TR924-S.1*

Table S.1
Phases of the Translation of CER into Clinical Practice

Phase	Description
Generation	Generation includes the design and conduct of the CER study; it involves primarily funders and CER researchers, but research priorities are influenced by the needs of multiple stakeholders, including scientists, the public, and policymakers.
Interpretation	Stakeholders ascribe meaning to CER results based on a number of factors, including the strength of evidence, applicability of the evidence to the potential adopter's practice setting, personal experience, and messages received by other stakeholders (e.g., professional societies, industry, media, and opinion leaders).
Formalization	Formalization is the process by which the interpretations of CER results are converted into guidance instruments such as clinical-practice guidelines, performance measures, and quality improvement tools. Multiple stakeholders may play roles in formalization through participation in guidelines committees, regulatory committees, and performance-measure development and endorsement processes.
Dissemination	Dissemination is the process by which CER information and/or associated tools designed to influence practice is actively transmitted to stakeholders. It typically promotes (or discourages) implementation of a new practice but may also have the goal of promoting a particular interpretation of the CER results.
Implementation	Implementation is the adoption of new clinical practices based on CER results. Implementation decisions may depend on a wide range of factors, including the dissemination of messages and the successful embedding of CER-related clinical guidance into tools that facilitate practice change, as well as the local market, regulatory, and professional context that may promote or impede changes. The implementation phase takes place primarily in local practice contexts.

Table S.2
Case-Study Topics

Topic	Type of Comparison	Reference
Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE): atypical antipsychotic drugs versus conventional antipsychotic drugs for schizophrenia	Medications	Lieberman, Stroup, et al., 2005
Clinical Outcomes Utilizing Revascularization and Aggressive Drug Evaluation (COURAGE): percutaneous coronary intervention versus optimal medical therapy for chronic stable angina	Medication versus procedure	Boden, 2007
Spine Patient Outcomes Research Trial (SPORT): surgical versus nonsurgical treatment for lumbar spinal stenosis	Procedures	Weinstein, Tosteson, et al., 2008
Comparison of Medical Therapy, Pacing, and Defibrillation in Heart Failure (COMPANION): optimal medical therapy versus cardiac resynchronization therapy versus combined cardiac resynchronization therapy and defibrillator therapy for patients with moderate to severe heart failure	Procedures	Bristow, Saxon, et al., 2004
Computerized physician order entry (CPOE): interventions to prevent serious medication errors	Delivery-system interventions	Bates et al., 1998

After selecting the case-study topics, we examined the peer-reviewed literature to obtain information on the extent to which the CER evidence led to practice change, the key stakeholders engaged in translation, and the specific dissemination activities involved. We developed a preliminary list of potential discussants for each topic and extended invitations to a set of discussants who, taken together, could provide perspective on all phases of the CER translation process. We developed a core discussion guide that would enable us to address key topics in a consistent manner across case studies and conducted 53 discussions with individuals or groups, including researchers involved with the CER studies, practicing physicians, leaders of professional societies, representatives of funding agencies, patient advocates, decision support developers, directors of quality improvement organizations, senior executives of health plans, Medicaid directors, journal editors, and leaders of integrated health systems.

CATIE Case-Study Summary

Background

The National Institute of Mental Health (NIMH) funded the $42.6 million CATIE study in 1999 to compare the effectiveness of a first-generation antipsychotic (perphenazine) to three second-generation medications: olanzapine, quetiapine, and risperidone. CATIE was considered a landmark trial because of its size, duration, and public sponsorship. Furthermore, it was designed to be generalizable to real-world clinical settings by using limited exclusion criteria, enrolling patients from diverse settings, and permitting flexible dosing protocols. Prior to CATIE, the optimal choice of drug treatments for patients with schizophrenia was disputed for at least four reasons: (1) uncertainties surrounding effectiveness in controlling psychotic symptoms; (2) uncertainty about the relative incidence of side effects, including tardive dyskinesia and metabolic side effects; (3) some evidence that second-generation antipsychotics improve cognition; and (4) the costliness of second-generation antipsychotics. These concerns prompted many to call for a rigorous assessment of the overall value of the second-generation antipsychotics.

Results

The initial results from CATIE were released in 2005, and they surprised many. The trial found that perphenazine was as effective as olanzapine in terms of time to discontinuation of medication for any cause, the trial's primary outcome. Patients randomized to olanzapine also had the longest time to discontinuation of any group because of lack of efficacy, the largest weight gain, and significant increases in other variables associated with the metabolic syndrome. The CATIE investigators concluded that perphenazine could not be rejected as an inferior treatment.

Lessons Learned

The CATIE case study highlighted the role of pharmaceutical manufacturers in shaping and reinforcing beliefs about the relative superiority of second-generation antipsychotics, both directly (through marketing and detailing) and indirectly (through key thought leaders) well in advance of the conduct of a CER study. By the time the CATIE results were released, these efforts had cemented beliefs about the various classes of antipsychotics. Practice patterns do not appear to have changed in the five years following the publication of the trial's results. Professional societies did not strongly advocate for practice changes based on the results.

Guidelines eventually changed but had limited impact. Performance measures were not updated to reflect the trial's findings and may have continued to reinforce existing prescribing patterns. Professional societies and advocacy organizations challenged the results of the trial in an effort to protect provider autonomy and preserve access to medications, respectively. Public payers were initially unwilling to enact policies that might limit the treatment options of patients with schizophrenia, given the relative lack of access to care for this population and the potential backlash from advocacy organizations.

A number of strategies might be used in CER trials that share some of the characteristics of CATIE to promote the uptake of results into practice. With regard to CER generation, methodological choices (in CATIE, the exclusion of patients with tardive dyskinesia from the perphenazine group) may limit the perceived generalizability of the findings and cause physicians to distrust the results. In the CATIE case, providers had strong prior negative experiences involving adverse outcomes such as tardive dyskinesia. Failure to design the trial to address one of the main beliefs driving use of a treatment—e.g., that second-generation medications were safer with respect to the incidence of tardive dyskinesia—meant that an issue important to prescribing providers might be perceived as being inadequately addressed by the study.

Interpretation and formalization faced formidable but predictable difficulties given the strength of established beliefs about the superior efficacy and safety of second-generation antipsychotics. Indeed, it proved quite difficult to change the deeply ingrained belief system founded on industry-funded studies. Interestingly, critiques of the study methodology that did not address harms may not have significantly influenced prescribing practices. Likewise, it should be borne in mind that professional societies can be expected to generate guidelines reflecting their professional interests if the study results leave room for such interpretations by failing to produce a "clear winner." Timely updates to quality measures that reflect the new CER evidence may prove critical to motivating early changes in practice.

Dissemination and implementation strategies should be vigorous and multipronged. In the case of CATIE, academic detailing within closed systems eventually proved effective in some cases, but early efforts were constrained by doubts about the value of the findings, and even a clinical decision support prompt faced initial resistance. A key element to success was presenting physicians with their actual practice data, which often showed how far they diverged from the ideal. Finally, future adverse-event surveillance systems (or registries) may help to resolve lingering questions about the relative side-effect risks of alternative antipsychotics, such as tardive dyskinesia, relative to cardiovascular disease.

COURAGE Case-Study Summary

Background

The COURAGE trial compared the risk of cardiovascular events among patients with stable coronary artery disease (CAD) assigned to a treatment strategy of intensive pharmacologic therapy and lifestyle intervention (optimal medical therapy) alone and patients who were assigned treatment with percutaneous coronary intervention (PCI) followed by optimal medical therapy. No study before COURAGE had included the intensity of medical therapy attempted in the trial, which included the use of aspirin, beta-blockers, angiotensin-converting enzyme (ACE) inhibitors, statins, and clopidogrel, as well as diet, exercise, and smoking-cessation counseling. Medication doses were repeatedly intensified in pursuit of aggressive blood-pressure and LDL-

cholesterol targets. The fundamental question addressed by COURAGE was whether the use of PCI to reverse narrowing of the coronary arteries in conjunction with optimal medical therapy would provide patients with stable CAD a greater reduction in the risk of myocardial infarction (MI) and death than medical therapy alone. While PCI has been shown to provide substantial benefit for patients being treated for emergency conditions, the benefit of the procedure in the stable CAD population had not been conclusively demonstrated. In the years preceding COURAGE, most clinical trials that assessed these end points were small and underpowered.

Results

The COURAGE trial found that as an initial management strategy in patients with stable CAD, PCI did not reduce the risk of death, MI, or other major cardiovascular events when added to optimal medical therapy. The findings reinforced existing practice guidelines, which stated that PCI can be safely deferred in patients with stable CAD, even in those with extensive, multivessel involvement and inducible ischemia, provided intensive, multifaceted medical therapy is instituted and maintained.

Lessons Learned

Both our discussions and a high-quality empirical analysis indicate that the COURAGE trial did not have an impact on clinical practice. The trial may have had an important indirect effect on practice by encouraging the integration of appropriateness criteria for coronary revascularization into decision support tools and into data collection for registries. How much these efforts will facilitate practice change remains unclear. Efforts to use appropriateness criteria in quality improvement are nascent, and while they have yet to be used in an accountability or payment context, there is increasing interest in them among policymakers. These initiatives will be most effective once reimbursement systems create demand for them. Changes in the organization of cardiology practices, driven in part by the movement toward accountable-care-organization (ACO) payment models, may be the single most important determinant of the future adoption of findings from COURAGE and other CER evidence.

Several strategies may improve uptake for CER trials that share some of the characteristics of COURAGE. In the generation phase, research should focus on a decision point sufficiently upstream to meaningfully impact decisionmaking. A critical driver of the use of PCI is the initial decision to refer a patient to an interventionist, since this tends to create an expectation that angiography and PCI will follow. The COURAGE trial did not address the initial referral decision directly. Rather, it addressed a decision point later in the pathway to PCI—after patients have undergone angiography—at which the utility of decision support and patient decisionmaking aids may be suboptimal. Current and proposed trials are focusing on decisions that occur prior to angiography, and these may have a greater impact on clinical practice. Other design problems to avoid include the potential for significant patient crossover or excessive time to complete the study. However, discussions with stakeholders suggest that criticisms of the trial design probably had only a minor influence on practice patterns post-COURAGE.

Interpretation and formalization can languish if study findings confirm current guidelines, even if they contradict current practice. Prior to COURAGE, practice guidelines were based on very weak evidence, promoting physicians' inclination to disregard them, but since the COURAGE results reinforced the guidelines, there was less impetus to revise them. A CER result that necessitates a change in guidelines may have more impact. Similarly, unless payers

and other stakeholders have the ability to collect relevant appropriateness data, they will have no incentive to develop reimbursement policies based on guidelines or appropriateness criteria.

Dissemination and implementation may be either advanced or retarded by several factors, but in this case, psychological aspects appear vital. For example, while registries may have influenced practice (by incorporating performance measures and appropriateness criteria into their design), their influence to date on appropriate use of elective PCI appears modest. Similarly, payer limits on upstream diagnostic procedures may have somewhat dampened demand for PCI, as might accountable-care reimbursement schemes in the future. Psychological factors, including concerns about harm and physician response to popular media coverage regarding PCI overuse, may have more significantly modulated the tendency to intervene aggressively. However, strong financial and psychological factors still incline both providers and patients to favor PCI. As one discussant put it, even without financial incentives, "interventionalists love to intervene." By all accounts, both clinicians and patients may underestimate the effectiveness of optimal medical therapy, and patients may not be informed of, fully understand, or seek out available information on the benefits and risks of PCI. Patient decision aids may play a key role in helping to address this information gap. However, to be maximally effective, such decision aids will have to be implemented in settings where financial incentives do not promote PCI and before patients have progressed to the point where intervention becomes inevitable.

SPORT Case-Study Summary

Background

The principal clinical question motivating SPORT was whether surgical treatment options were superior to nonsurgical treatment for patients with low back pain related to lumbar spinal disorders (disc herniation, spinal stenosis, and degenerative spondylolisthesis). Our case study focused only on the subpopulation with spinal stenosis. Prior to SPORT, the Maine Lumbar Spine Study, a prospective cohort study enrolling 148 patients, was the largest study comparing the effectiveness of alternative treatments for spinal stenosis, and it found that surgical patients had better outcomes than patients receiving nonsurgical treatments. A 2005 Cochrane review summarizing the evidence prior to 2000 suggested that the relative efficacy of surgery was not established, because existing trials were small and enrolled patients both with and without degenerative spondylolisthesis. A large, randomized CER trial was considered necessary to provide stronger evidence on the effectiveness of surgical treatment for spinal stenosis among patients who did not have degenerative spondylolisthesis.

Results

SPORT's intention-to-treat analysis showed that surgery was more effective than non-operative treatment on the SF-36 bodily pain scale and on patients' self-reported ratings of symptom improvement but on few other primary or secondary outcomes. However, patients with spinal stenosis had very high rates of crossover after randomization (as was the case for the subpopulations with disc herniation and degenerative spondylolisthesis). Only 67 percent of patients randomized to the surgical arm underwent surgery, while 43 percent of the patients randomized to nonsurgical treatment underwent surgery within two years of the baseline assessment. For this reason, the data from the randomized cohort were combined with data from an independent and concurrent cohort study and analyzed as a single observational study

comparing patients who underwent surgery with those who did not (an "as-treated" analysis). The observational analysis found that surgery was superior to nonsurgical treatment across all primary and secondary outcomes, and the advantage was sustained over two years of follow-up.

Lessons Learned

SPORT appears to have had little impact on clinical practice, and the seeds of its low impact appear to have been sown primarily in the generation phase. The study design, which was unblinded, also allowed for very large patient crossover. As a result, what was intended to be a randomized controlled trial (RCT) with an "intention-to-treat" analysis had to also be treated as an observational cohort study using an as-treated analysis. Analogous studies have avoided these difficulties, suggesting that they are not inherent in this type of CER but can be forestalled by careful study design and execution.

In the interpretation phase, the RCT results suggesting limited benefits from surgery were discounted because of high rates of patient crossover. In contrast, the observational cohort study, at least within the spinal surgery community, confirmed the relative advantage of surgery, which was already the prevailing method of treatment. Interpretation was further complicated by the study's lack of detail on subgroups, which made it hard to determine whom surgery would benefit most, as well as the (possibly erroneous) perception that the surgical techniques used in the study were already outdated. While presenting competing analyses may have opened the results to conflicting interpretation, the observational results alone produced different interpretations regarding the magnitude of the benefit provided by surgery.

The SPORT case study also highlights the challenges in weighing the relative strengths and weaknesses of RCTs and observational cohort studies and the selective use of evidence during the formalization phase. Multiple specialty societies, possibly influenced by various levels of industry sponsorship, issued competing and conflicting guidelines, while relevant data from European studies were generally discounted or ignored. Registries might help to bolster guidelines or generate appropriateness criteria, but since effectiveness outcomes from spinal surgery are often subjective, registries may be best suited to report on harms. Registry penetration appears quite low in orthopedic surgery, and financial incentives are not aligned to promote participation by surgeons.

While dissemination of the SPORT results appeared to be far-reaching, messaging about them emphasized the benefits of surgery rather than the significant clinical improvement among patients in the nonsurgical group and the relatively small difference in clinical benefit between the groups. Referring providers appear to be the optimal point for dissemination of the results, since referral to a surgeon is usually followed by surgery. Intense marketing of spinal hardware by the device industry may override the results of clinical trials, and, as SPORT illustrates, messages may be vague and selective, omitting key evidence provided by trials. Similarly, payers and purchasers, faced with both the "positive" results from the observational cohort analysis and the "equivalence" results from the intention-to-treat analysis, appear to have accepted the primacy of the observational cohort analyses and did not enact policies restricting the use of decompression surgery. However, there are now some early examples of more nuanced and data-driven reimbursement policies focusing on related procedures (e.g., fusion surgery).

Nevertheless, at the implementation phase, strong financial incentives favor surgical over nonsurgical treatment. The alignment of financial incentives among physicians, hospitals, and device manufacturers appears to have increased the use of complex procedures despite

uncertainty about their effectiveness and considerable evidence of greater risks. Countering this trend are radiology benefits managers (RBMs), which may reduce inappropriate upstream diagnostic procedures, and a potential future role for patient decision aids. While the SPORT results can be viewed as both flawed and confirmatory of current practice, the trial was successful in providing quality data on the relative risks and benefits of surgery, and these data have been integrated into patient decision aids. Those tools might ultimately change clinical practice by more fully incorporating patient preferences into decisions about surgery. Currently, few incentives encourage the use of such shared decisionmaking or a more rigorous informed-consent process. The use of these techniques early in the pathway leading to surgery will be critical to their overall effectiveness. Incentives to promote the spread of patient decision aids and efforts to improve the appropriate use of diagnostic imaging represent the most important strategies for changing clinical practice in the future.

COMPANION Case-Study Summary

Background

Cardiac resynchronization therapy (CRT) can improve the health status of patients with heart failure (HF) by electrically stimulating the heart to improve synchronization of pumping. Most HF patients appear to be at high risk for potentially fatal derangements in the heart's electrical activity. Implantable cardioverter defibrillators (ICDs) protect against sudden death from abnormal rhythms but do not reduce HF symptoms. The principal question addressed by the COMPANION trial was whether adding CRT alone or combined with ICD treatment (CRT versus CRT-D) to the medical management of HF patients with conduction abnormalities not only improved functional measures but also reduced hospitalizations and all-cause mortality (Bristow et al., 2004). Previous studies did not have sufficient power to detect a survival advantage from combined therapy.

Results

The COMPANION trial showed that patients assigned to the CRT group and the CRT-D group both had a statistically significant improvement of 17 percent over medical therapy alone in the combined end point of death and hospitalization from any cause. While adding an ICD to CRT did not appear to benefit patients more than CRT alone, it did show a trend toward reducing 12-month all-cause mortality (12 percent versus 15 percent). The results of this study imply a clear survival and quality-of-life benefit from adding CRT (with or without CRT-D) to optimal medical therapy for patients who suffer from HF with delayed ventricular conduction. This contrasted with the current practice at the time, which was to use CRT for HF patients but withhold ICD devices given both safety concerns and a lack of proven benefit.

Lessons Learned

Uptake of the CER results following publication of the COMPANION study has been uneven. Recent estimates indicate that there is both significant underuse of CRT among potentially eligible HF patients and also fairly frequent CRT-D use in patients who lack an indication for it.

In contrast to the other CER case studies, COMPANION generated relatively few controversies in the generation and interpretation phases. The results were fairly readily accepted, the main disputes being over the degree to which they could be generalized to HF patients who did not meet the original inclusion criteria. Formalization of the COMPANION

results was relatively rapid: specialty-society guidelines were updated promptly, which promoted their uptake, at least among proceduralists and HF-management specialists. No primary-care specialty-society guidelines were issued. In addition, the specialty-society guidelines left open the appropriateness of CRT-D. This and other factors would in turn contribute to an ineffective dissemination phase.

This case study illustrates the critical role dissemination plays in translating CER research into practice. Essentially all dissemination activities focused on interventional cardiologists and HF specialists rather than referring physicians. Specialty societies, industry, and other continuing-medical-education (CME) producers all directed their educational efforts toward those groups. Most primary-care providers (who manage many HF patients) remain unaware of the COMPANION results. In addition, those primary-care providers and general cardiologists who took an interest in the study findings were confronted with conflicting and ambiguous guidelines. This generated considerable confusion and a reported reluctance to refer patients for CRT. Future similar CER dissemination should focus significant effort on providers further upstream in the decision pathway and on delivering clear, unambiguous referral criteria. However, the COMPANION case is not only a cautionary tale. HF registries have had a significant positive impact through publication of high-profile studies illustrating inappropriate ICD use. Similarly, recent limited experience with clinical decision support tools shows that they can be very effective in prompting appropriate referrals and discouraging inappropriate procedures, but such tools for CRT or CRT-D appear to be rare.

In the implementation phase, imprecise guidelines and evidence-neutral reimbursement policy may contribute to the use of CRT-D for inappropriate indications. Reimbursement policies, particularly those of Medicare, significantly favor CRT-D implantation over CRT alone, despite evidence that adding the ICD has a very high marginal cost relative to the benefits it confers. As with other studies, referral to an interventionist is also tantamount to ordering the procedure. This tendency is compounded by open guidelines that allow the CER results to be cited as justifying use in patients who would not meet study inclusion criteria. While primary-care physicians and some general cardiologists fail to refer many potentially eligible patients, dedicated HF clinics have been much more successful in achieving appropriate referrals, as well as avoiding inappropriate ones. Such clinics may serve as a model for implementing analogous CER results. Currently, patients are not generally equipped to participate as fully informed partners in the clinical decision, and decision aids are not readily available, but it is likely that such decision aids could significantly improve appropriate use of CRT if physicians were given stronger incentives to use them.

CPOE Case-Study Summary

Background

During the 1990s, experts debated the optimal approach to reducing medication errors. Some were not persuaded that traditional paper-based ordering systems were a significant problem or that computer-based ordering systems alone (e.g., for medications and lab tests) would be more effective in reducing the rate of medication errors than nurse-focused, pharmacist-focused, or team-based interventions. The principal CER question leading up to Bates's 1998 study was whether computerized physician order entry (CPOE) could reduce medication errors and medication-related adverse events among hospitalized patients more effectively than other interventions. To address this question, Bates and colleagues compared the effectiveness of

CPOE alone with CPOE plus a team intervention for reducing the number of unintercepted serious medication errors. The study, conducted within six units at Brigham and Women's Hospital (Boston, Mass.), used a pre/post design, where the rates of medication errors prior to CPOE adoption were compared with the rates during the ten months following CPOE adoption. The CPOE system allowed physicians to select from a menu of medications defined by the hospital formulary, with default dose and dose ranges provided for each medication, as well as automatic checking for common drug allergies and drug-drug interactions. The team intervention centered on pharmacy-specific process changes, including changing the role of the pharmacist, standardizing labeling of intravenous bags, and implementing a pharmacy communication log so that the nursing staff could better communicate with pharmacy staff.

Results

The analysis indicated that the team intervention had no incremental benefit over the implementation of CPOE alone, so the intervention arms were pooled. Between pre and post periods in the same hospital units, unintercepted serious medication errors (the study's primary end point) decreased by 55 percent. Unintercepted potential adverse drug events (ADEs)—a secondary end point—declined by 84 percent. The authors of the study concluded that, on the basis of these results, other hospitals should consider CPOE adoption as the principal intervention to reduce unintercepted serious medication errors.

Lessons Learned

In contrast to the adoption of a new medication or device that does not require significant changes to the work of clinicians and staff but instead funnels through the existing workflow, the adoption of quality improvement strategies faces serious barriers because they may require significant changes to organization, financing, and staff work. These barriers may neutralize the impact of even outstanding CER evidence. While we selected CPOE as an example of a delivery-system intervention for which there is published CER evidence, it is worth noting that although CPOE has features that are typical of many such interventions, it also has distinct features that may be easier to implement. The CPOE case study suggests a number of lessons about translating delivery-system-related CER into new practices.

First, CPOE is both a new technology and a new set of workflow requirements. These features of the intervention are complex and require substantial up-front investment, as well as coordination, communication, and long-run commitments from numerous stakeholders with potentially conflicting goals. The staff involved in implementation of CPOE must make nontrivial changes in workflow; like other technology-based delivery-system interventions, CPOE requires dramatic changes in individual process and social interactions with peers.

Second, CPOE is a variable technology with evolving features and functionalities. It has numerous meanings across a wide range of hospitals and different vendors, depending on their needs and existing health information technology (HIT) capabilities. This poses challenges for end users (particularly hospital executives) who wish to use CER evidence for decisionmaking. Our case study suggests that these individuals often struggle to conceptualize the intervention and consequently may find it difficult to assess the applicability of the results to their own settings.

Third, the financial investment in CPOE is substantial, and key leaders must have clear reasons and plans for implementation to overcome resistance from staff. Successful CPOE implementation appears to require financial incentives to improve the business case, and the

experience of early adopters suggests that organizational factors and missions can be significant enablers even when financial incentives are not aligned.

Fourth, the target stakeholders for CER results concerning CPOE are more diverse than the typical users of other types of CER studies. They include hospital executives and technology vendors, in addition to physicians, pharmacists, and other clinical staff. This may increase the complexity of messaging to achieve effective and consistent dissemination of the CER results.

Despite the unique features of CPOE, similar delivery-system interventions based on CER evidence (particularly those that improve patient safety) may also benefit from some combination of strong mandates, systematic standards, and financial incentives that improve the business case for implementation.

Root Causes of CER Failure to Rapidly Change Clinical Practice

A myriad of factors influence whether CER is successfully translated into clinical practice. However, our synthesis suggests that some of these factors are "root causes" in the sense that they are fundamental and may represent high-leverage points for action to improve adoption of a new practice. We identified five root causes of failure when CER is slow to change clinical practice. These root causes manifest themselves in somewhat different ways across the case studies, appear to explain the strategies of the many stakeholders with an interest in CER, and typically exert their effects over multiple phases of the CER translation process.

1. ***Financial incentives are primary drivers of adoption of new clinical practices whether or not the practices are supported by CER evidence. CER results that threaten the financial interests of a stakeholder will be challenged at all phases of the CER translation process.***

The most fundamental determinant of successful CER translation is the extent to which the economics of adopting a new clinical practice are favorable to providers and patients. Our case studies on the comparative effectiveness of interventional and noninterventional procedures highlight the perverse consequences of fee-for-service reimbursement as a driver of the use of procedures that CER evidence shows have little or no marginal benefit. Once patients are referred to interventional specialists, even if only for consultation, there is a high likelihood that they will receive an invasive procedure.

Our case studies highlight the role of financial incentives in influencing more than only the implementation phase of the CER translation process. In particular, financial incentives may supersede CER evidence in influencing the adoption of new clinical practices in the following ways:

- Stakeholders with a financial interest in the outcome of a CER study may seek to influence its design in order to increase the odds that its results will favor them, or they may initiate efforts to critique and thus undermine potentially unfavorable CER studies at the time the studies are enrolling participants. Critiques of a CER study design by interested stakeholders may peak when the results are released to maximize the likelihood that the study will be viewed as methodologically weak.
- The interpretation of CER results through a dynamic scientific debate among stakeholders appears to be influenced by financial incentives of the participants.
- The formalization of guidelines and measures based on CER evidence may be influenced in subtle ways by financing, and professionals have few financial incentives to facilitate

the development of performance measures unless they will be paid based on the measure results. If guidelines do not evolve with the CER evidence, other formalization activities such as the modification of quality measures will be delayed or simply fail to occur.

- The dissemination of new practices is expensive, and the lack of financing for dissemination activities to support CER-based practices may be an important impediment to change. Aggressive dissemination activities directed toward payers may cause them to focus narrowly on areas where practice variation is extensive, where evidence clearly does not support a practice, and where risks to patients are unambiguous.
- Physicians working within a larger organizational context may be more likely to use performance measurement and feedback, patient decision aids, clinical decision support tools, and registries, all of which have the potential to increase responsiveness to CER.

Despite the seemingly powerful influence of financial incentives favoring both the status quo and an accelerating panoply of new procedures, recent trends, including the emergence of innovative payment models and new types of physician organizations, provide some basis for optimism that CER evidence can be more influential in the future. In addition, activities that curtail the influence of financial interests in each of the phases of CER translation (such as the recent Institute of Medicine [IOM] report calling for greater transparency and integrity in the guideline development process) may reduce the countervailing forces that work to undermine or neutralize even the best CER evidence. Payers are also actively engaged in horizon-scanning for CER evidence that may form the basis of policies before practices become widespread in the community.

2. ***Even the best CER studies may fail to produce an unambiguous "winner," so it may be difficult to achieve a consensus interpretation of the results.***

CER studies that produce clear "winners" (i.e., showing unambiguously that one treatment is better than another or that two treatments have essentially equivalent effectiveness) should be more likely to change practice because they are difficult to challenge. However, our case studies suggest that even among the best-designed and -conducted CER studies, unambiguous outcomes are likely to be rare. Many factors increase the risk of an ambiguous result from a CER study, including design factors (e.g., use of active comparison groups rather than placebos), differential weighting of end points by stakeholders, and differences in provider equipoise for recommending treatments. Persuading stakeholders about treatment equivalence may be much more difficult than persuading them of treatment superiority. CER studies that produce ambiguous results open the door to selective interpretation, may undermine consensus interpretation of the results, and may fail to promote guideline updates by professional societies or the formation of coverage policies by payers. In cases where one treatment is found to be unambiguously harmful (e.g., the Women's Health Initiative), clinical practice has been known to change rapidly, and our findings confirm to some extent that general rule, based on anecdotal reports suggesting decreased use of olanzapine in the post-CATIE period. Adverse-event data from registries may help to identify more unambiguous "losers" over time.

While ambiguity may lead to incomplete use of CER results and may limit the potentially attainable change in clinical practice, the lack of "winners" does not invariably mean that the CER fails to have an impact on clinical practice. Many discussants indicated that the goal of CER is not identifying "winners," but generating information to help physicians and patients arrive at satisfactory treatment decisions. Several of our case studies might have reassured

physicians and patients that moderate-dose conventional antipsychotics or less-aggressive therapies can have benefit comparable to that of more-aggressive therapies.

3. ***Cognitive biases play an important role in stakeholder interpretation of CER evidence and may be a formidable barrier to clinical-practice change.***

At least three cognitive biases may influence the way in which physicians and other stakeholders interpret new CER evidence. First, confirmation bias, the tendency for a stakeholder to embrace evidence that confirms preconceived notions of treatment effectiveness and reject evidence to the contrary (while typically criticizing the studies on methodological grounds), may reinforce established practice patterns. Stronger study designs (emphasizing particularly the generalizability of findings) and careful monitoring of study conduct (particularly to prevent crossover for randomized study designs) may preempt these critiques and counteract the influence of confirmation bias. A second bias is the belief that intervening aggressively is better than inaction, even when the marginal benefit is small. This bias may be reinforced by perverse financial incentives and by providers' perceived risk of malpractice liability if they fail to act; however, more complete data on treatment harms or heterogeneity of benefits may promote greater equipoise among both physicians and patients. A third cognitive bias, which may be reinforced through messaging by interested stakeholders, is the tendency to perceive new technologies as superior to older technologies—a problem with lengthy CER studies, during which technology often advances. Adaptive study designs could provide the flexibility to allow the evolution of treatments through the course of a trial, but these approaches were not used in any of the trials we studied, with the exception of CATIE.

Strategies for mitigating cognitive biases are available, but their effectiveness is not completely clear. Enhancing the transparency of stakeholder positions by using approaches that foster explicit formal decisionmaking processes is one approach to mitigating cognitive biases in a policymaking context. Disclosure of financial and intellectual conflicts of interest is another strategy used by the IOM and others. Regulation of detailing and direct-to-consumer advertising may also be effective.

4. ***The questions posed by a CER study and its design may not adequately address the needs of end users or focus adequately on the clinical decisionmaking opportunities that have the greatest potential to influence clinical practice.***

Our case studies suggest that CER faces potentially unrealistic expectations on the part of multiple end users. First, there is an unavoidable tension in the design of CER studies between supporting personalized medicine and supporting clinical policymaking, which requires generalizable results on larger populations. Second, as demonstrated in CATIE, CER studies may not be designed with a comprehensive or explicit understanding of the beliefs and concerns of clinical practitioners, such as their preoccupation with the relative safety of classes of antipsychotics rather than their relative effectiveness. Third, head-to-head comparisons of treatments may help providers select appropriate treatments in the later stages of clinical decision algorithms, but CER concerned with upstream diagnostic tests or procedures may have a larger impact on patient outcomes and the overall value of care. If distinct providers (e.g., primary-care providers) are responsible for upstream decisions to refer, and if both providers and patients face weaker incentives to choose an intervention (compared to interventionists), their decisionmaking may be more readily influenced by evidence.

5. ***Clinical decision support and patient decision aids can help to align clinical practice with CER evidence, but they are not widely used.***

Perverse financial incentives and lack of accountability for implementation have limited the production and dissemination of both clinical decision support tools and shared-decisionmaking aids. Decision support tools to promote evidence-based diagnostic testing and appropriate referral to specialists are uncommon, although both may lead to the use of treatments that are better aligned with CER evidence. Well-informed patients who make treatment decisions according to their preferences may ultimately serve as a counterweight to providers who lack equipoise. However, the prevalence of direct-to-consumer advertising has grown, and this may create or reinforce misconceptions about treatments. Even if incentives for adopting these tools were better aligned, the challenge of integrating them seamlessly into clinical practice has not been solved, and limited HIT infrastructure and inadequate provider training on shared decisionmaking may continue to pose barriers to the implementation of these tools in the coming years.

Limitations of This Study

Because our primary objective was to identify and synthesize themes across case studies, we struck a specific balance between depth of content and breadth of inclusion of case studies. Our sample of expert discussions and stakeholder perspectives for each of the case studies was limited, and a larger sample of end users of CER from a diversity of practice settings could have identified additional barriers and enablers. The perspectives of the device industry and, to a lesser extent, the pharmaceutical industry are also underrepresented, despite considerable outreach efforts. Most potential discussants declined to participate, citing concerns about the sensitivity of the information they might be asked to share.

We chose not to use a formal qualitative research methodology that included the coding of themes with the use of specialized analytic software. Biased interpretations of the data by the research team were mitigated by requiring a minimum of three investigators to be present for each discussion, and we used email follow-up for areas that needed clarification. Our root-cause analysis drew primarily on themes that were mentioned repeatedly by stakeholders. Because of the limited scope of topics and the limited number of discussions we were able to hold, some of our findings regarding the root causes of failed CER translation or the facilitators of practice change may not be generalizable to other topics or to a broader range of practice settings.

Policy Implications

While the root causes of failure to translate CER evidence into clinical practice are formidable, they are not insurmountable. After reflecting on the policy implications of our case studies, we identified a range of policy options that can address the root causes and promote more effective translation of CER evidence into clinical practice. Each of these policy options can be categorized into one of the following domains: governance, standards, financing, professionalism, marketing and education, and research and evaluation. Each of the policy options would be optimally deployed within a healthcare system having a CER-enabling infrastructure that

1. Enables generation of CER that is more relevant to decisionmakers

2. Enables more-effective translation of research into practice
3. Enables more-effective evaluation of the impact of translation activities.

The policy options presented here are not intended to create a centralized command-and-control infrastructure that determines the CER agenda. Changes in the translation process, such as reengineering financial incentives, must be carried out by a diverse set of public and private stakeholders, and the changes must address a remarkable diversity of payment arrangements. The policy options we suggest could bring greater coherence and transparency to the process of CER translation, achieve greater balance of the influence of stakeholders participating in CER translation, and enhance the voice of the public and patients whose health outcomes depend on effective, safe, and affordable care. Enacting some or all of these options could be expected over time to modify the financial and other incentives that shape clinical decisionmaking so that decisions will be increasingly based on evidence rather than other considerations.

Governance

Create a transparent governance mechanism with oversight of the CER translation process:

1. *Include patient or consumer representatives.* Patients are ultimately the financers of CER and the key beneficiaries of the clinical practices guided by the research. They should be engaged to ensure that the CER translation process is well informed by the end-user perspective and to provide a counterbalance to other stakeholder interests.
2. *Include public and private payer and purchaser representatives.* As stewards of the financing of healthcare and representatives of their member or customer interests, payers and purchasers may be able to identify the specific opportunities for high-value care that would be especially amenable to CER and for which modification of payment or coverage policies could be especially effective in optimizing clinical practice.
3. *Enable and support public comment opportunities.* Vigorous solicitation of public comment with verification and full disclosure of the potential conflicts of interest (both financial and intellectual) of those who offer comments can enhance the credibility of the CER enterprise.
4. *Institute strong policies on disclosure and management of potential conflicts of interest.* CER evidence that is perceived as biased by financial interests will lose credibility and can impede the take-up of new practices; it might also be countered by policies that identify and manage potential financial, institutional, and intellectual conflicts of interest.
5. *Use the governance mechanism to generate a prospective public record of stakeholder expectations.* Documenting the positions of relevant stakeholders at the outset of CER studies with respect to the study objectives and the parameters around which the results should be interpreted creates a public record of expectations of each stakeholder and may discourage post-hoc efforts to undermine the credibility of studies that produce results contrary to the interests of specific stakeholders.

Standards

Support and enhance creation of standards for CER generation and translation:

1. *Incorporate data elements that are critical to translation activities into the proposed national CER registry.* Creating explicit standards for the description of CER study objectives, design, sampling, and causes of heterogeneity in sampled populations within a CER registry may be useful to guide formulation of a consensus interpretation of results and to avoid post-hoc reframing of the questions and implications by stakeholders to serve their interests.
2. *Encourage development of standardized electronic clinical data systems (clinical registries).* Standardized electronic clinical registries can enable rapid and low-cost CER, can provide data for longitudinal tracking systems to evaluate the impact of CER translation activities on clinical practice patterns, and may better support decisionmaking because clinicians and patients perceive data derived from them to be trustworthy.

Financing

Encourage public financing of CER translation and promote the use of CER evidence in payment programs:

1. *Provide direct and indirect support for formalization of CER evidence and the dissemination of CER-based clinical practices.* Public financing for both translation of high-profile CER evidence in guidelines, quality measures, and clinical decision support and subsequent dissemination activities could counteract the influence of industry in the process of formalizing CER evidence that might otherwise undermine the evidence or selectively promote less effective or more costly practices.
2. *Promote the use of CER-based clinical practices through payment policy and incentive programs directed toward providers and patients (e.g., value-based purchasing).* Encouraging "prudent purchasing strategies" for public-sector payers based on CER evidence can assure that the financial incentives of providers and patients in the delivery of healthcare are well aligned to support clinical practices based on CER evidence and discourage practices that are not evidence-based.

Professionalism

Supporting professional consensus across the phases of CER translation:

1. *Foster and support a broad vision of professionalism in the governance of CER translation.* Broadly constituted professional committees may be able to produce balanced, consensus interpretations of CER results and may resolve differences of opinion about interpretation of those results in a transparent manner. Multispecialty clinical registries and health-information exchange may counteract the tendency to focus on narrowly defined subspecialty interests.
2. *Include training for professionals on the role of cognitive biases in diagnostic and treatment decisionmaking.* Training in the nature, role, and impact of cognitive biases can enable professionals to recognize the circumstances under which decisions and clinical

recommendations are prone to cognitive biases and to employ specific techniques that modify the decisionmaking context to compensate for these biases.

Marketing and Education

Promote demand for CER-based clinical services through public education and marketing:

1. *Promote patient demand for CER-based clinical services through shared decisionmaking that includes formal patient decision aids.* Shared decisionmaking involving the use of formal decision aids is the most prominent approach for assuring that patients can receive the best available evidence about alternative tests and treatments in a usable form. Decision aids could counter the messages that promote suboptimal clinical practices.
2. *Support "social marketing" campaigns for high-profile CER results to counteract the effects of industry-sponsored detailing and direct-to-consumer advertising.* Marketing campaigns, including detailing to clinicians and direct-to-consumer advertising, are generally aimed at exploiting cognitive biases, and this may impede the uptake of CER-based clinical services. Social marketing—the application of marketing techniques to promote behavioral change—has the potential to increase awareness and demand for evidence-based healthcare services by promoting greater patient engagement in medical decisionmaking.

Research and Evaluation

1. *Support research to identify the gaps in clinical decisionmaking that are the highest-priority topics for end users of CER.* Research on end-user needs can help identify high-priority topics and increase the relevance of CER to payers, professionals, and the public while fostering the selection of approaches for disseminating CER results tailored to the expectations of key stakeholders.
2. *Promote integration of CER registries with clinical registries to support evaluation of the impact of CER studies and the factors associated with successful translation.* Prospective evaluation of the impact of CER could be strengthened if available data sources can provide valid and reliable estimates of current clinical-practice patterns, which may be found in clinical registries. Clinical registries that are capable of providing longitudinal data on patients may enable these more complex studies and increase the relevance of results for end users.
3. *Support projects that develop unbiased and efficient methods for formalization of CER results.* Methods for developing and refining guidelines, performance measures, and clinical decision support tools are still a work in progress. Support for research and demonstration projects that develop and study new methods for formalization could lead to more effective, efficient, and unbiased tools.
4. *Support projects that enhance the utility of CER results by demonstrating and evaluating models for the use of decision aids by clinicians and patients.* Creating more effective decision aids, training professionals to use them, and developing strategies for embedding them in routine practice have all proven challenging. Research and evaluation projects that lead to better decision aids and their more effective use could increase the impact of CER.

5. *Support research into the ways in which CER evidence is used by integrated delivery systems.* Integrated systems may have unique perspectives on which CER topics are likely to have the greatest impact on clinical practice, and their CER translation experience may be invaluable. Future studies might involve engaging these organizations to elicit best practices in CER translation and evaluating which strategies may be transferrable to nonintegrated delivery settings.

The federal government is making a sizable investment in CER in the hope that the results will influence the decisions of clinicians and patients, optimize the quality of care, and lower costs. Given these goals, attention to the root causes of ineffective translation of CER evidence into practice seems critical. The list of policy options we have outlined is not exhaustive, but we believe that these options may provide guidance to a broad set of policymakers concerned with the organization and financing of healthcare. Taken together, the options suggest a number of different paths forward. Exploring multiple approaches may be appropriate in view of the large number of factors that impede CER translation. As the ARRA-financed CER portfolio begins to produce new evidence, a number of opportunities will arise in the near term to experiment with these strategies.