

# Better Evidence Better Health Care

**Industry leaders and policymakers  
call for a concerted effort in comparative  
effectiveness research.**

BY LOUISE KERTESZ

**R**EPORTS DOCUMENTING THE RESOURCES WASTED on medical procedures of questionable effectiveness are prompting industry leaders and policymakers to address these inefficiencies in the nation's health care system with new urgency. In many cases, information about which treatments are most effective simply does not exist. Often, patients fail to receive what are known to be the best treatments for their conditions because the information is not adequately disseminated or understood. We know little about whether new technologies—often with high price tags—are better than existing and oftentimes lower cost or less invasive technologies.

While organizations such as the federal Agency for Healthcare Research and Quality (AHRQ) continue to undertake and publish comparative effectiveness studies (see sidebar, "AHRQ's Comparative Effectiveness Efforts"), industry and government leaders say a more concerted effort must be made to compare the effectiveness of medical procedures and to make this information readily available to caregivers, payers, and policymakers. There is also a growing consensus that the cost-effectiveness of medical treatments must be considered along with their clinical effectiveness.

In April 2007, America's Health Insurance Plans became one of the first organizations to call for establishing "a new national entity to evaluate and compare the safety, efficacy, and cost effectiveness of new and existing health care treatments and technologies, including prescription drugs and medical devices." As part of a national strategy to improve safety and

quality, AHIP said that work in "comparative effectiveness" should set "a national research agenda that addresses known gaps in evidence and makes communication regarding ongoing research studies a national priority."

AHIP also called for "accelerating efforts to give patients and their physicians the information they need to make value-based health care decisions" and urged "the adoption of best practices" in treating patients.

In recent months, other leading organizations, including the Institute of Medicine (IOM), the Congressional Budget Office (CBO), and the Medicare Payment Advisory Commission (MedPac), have called for accelerating work in comparative effectiveness. This includes "designation of a single entity...with authority, overarching responsibility, sustained resources, and adequate capacity to ensure production of credible, unbiased information about what is known and not known about clini-

ISTOCK

# Decisions



cal effectiveness,” according to a recent IOM report, “Knowing What Works in Health Care: A Roadmap for the Nation.” The report makes recommendations on how to improve work in comparative assessment that various organizations are undertaking (see sidebar, “IOM Recommendations for Comparative Effectiveness Research”).

Last December, the CBO released a report, conducted at the request of Senate Budget Committee Chairman Kent Conrad, D-N.D., and Senate Finance Committee Chairman Max Baucus, D-Mont., “highlighting the potential for comparative effectiveness research to improve health care outcomes and control health care costs,” according to a press release from the senators. The report “outlines options for and implications of designing a new private and/or public sector initiative on comparative effectiveness research....Incentive structures based on evidence from comparative effectiveness studies could lead to cost savings in the nation’s overall health care system,” the release said.

Conrad said, “CBO’s report confirms that enhanced comparative effectiveness research can be an important tool in

controlling health care costs....We must get more value and efficiency out of the health care dollars we are spending. Increased comparative effectiveness research should be part of the solution.”

Congress has shown a strong interest in legislation that would increase comparative effectiveness research, points out J. Michael McGinnis, M.D., senior scholar at the IOM. A bill passed the House last year that would have increased AHRQ funding “almost tenfold...and there is good support on the Senate side for expanding our capacity in comparative effectiveness research.”

In addition, the health care proposals of both major Democratic presidential candidates and, to some extent, that of Sen. McCain have provisions to increase our knowledge of what works best in health care so as to decrease wasteful spending. Whatever the precise focus of comparative effectiveness legislation, “we will probably see something pass within the next 12 months,” McGinnis predicts.

Indeed, according to the release from Sens. Conrad and Baucus in December, they planned to introduce legislation this

## AHRQ’s Comparative Effectiveness Efforts

There are three major components to the Agency for Healthcare Research and Quality’s Effective Health Care Program, funded by Section 1013 of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003. The MMA authorizes AHRQ “to conduct and support research with a focus on outcomes, comparative clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services.” The underpinning of AHRQ’s work was created in 1997, when the agency established 15 Evidence-Based Practice Centers to perform systematic reviews and synthesize existing evidence to promote evidence-based practice. The centers also conduct comparative effectiveness reviews, an important aspect of which is to identify research gaps and recommend studies and approaches to fill those gaps.

“It’s impressive to see what we don’t know about what treatments

work for common diseases,” says Jean Slutsky, director, Center for Outcomes and Evidence, at AHRQ. Recently published reports such as those from *JAMA* on treatments for back pain and AHRQ on treatments for localized prostate cancer demonstrate the need to “identify what we don’t know and use that knowledge to conduct studies in important research gaps.” The centers have produced reports on cancer; diabetes; digestive system conditions; heart and blood vessel conditions; mental health; and muscle, bone, and joint conditions.

AHRQ also has “a network of research institutions doing rapid-cycle research,” Slutsky says. AHRQ created the network in 2005 to generate new knowledge and has made research reports on new evidence and analytical tools available, which cover many of the same conditions as those studied by the Evidence-Based Practice Centers.

The third component of AHRQ’s work is translating the knowledge

and information developed by the Evidence-Based Practice Centers and AHRQ’s research network into usable formats. The John M. Eisenberg Clinical Decisions and Communications Science Center at the Oregon Health Sciences University, established in 2005, translates “complex scientific information into short, plain language materials that can be used to assess treatments, medications and technologies,” according to AHRQ’s Web site. The center develops information summaries for consumers, clinicians, and policymakers. “The patient guides and clinician guides are paired, so [those groups] are able to talk about the results together,” Slutsky says.

AHRQ’s Effective Health Care Program Web site, [www.effective-healthcare.ahrq.gov](http://www.effective-healthcare.ahrq.gov), provides a comprehensive view of the program, including reports and directions for those who wish to get involved in suggesting topics for research and commenting on draft reviews.

year to create a national comparative effectiveness research initiative.

### Adding a Necessary Component

In a multi-stakeholder collaborative convened by AHIP, representatives from Wellpoint, Aetna, Merck, Johnson and Johnson, AHRQ, and others have been “trying to develop a more explicit and transparent way for decision makers to judge the strength of a body of evidence,” explains Steven D. Pearson, M.D., a senior fellow at AHIP. Pearson also is founder and director of The Institute for Clinical and Economic Review (ICER), a program based at Harvard Medical School that conducts comparative effectiveness studies. ICER is supported by funding from a diverse range of organizations including AHIP, health plans, and pharmaceutical companies.

In its assessments, ICER uses the Evidence-Based Medicine matrix developed by the AHIP-convened collaborative. The matrix is a “joint judgment of the level of confidence provided

by the body of evidence and the magnitude of the net health benefit—the overall balance between benefits and harms.” ICER uses the Evidence-Based Medicine matrix to arrive at a unique Integrated Evidence Rating, which evaluates a medical procedure not only for its comparative clinical effectiveness but also for its economic value or cost effectiveness.

ICER “identified accurately the need for payers as well as decision makers to have a standardized method” of evaluating evidence and to have “a clear understanding of the state of the evidence,” says Jed Weissberg, M.D., associate executive director, Quality and Performance Improvement, The Permanente Federation. But, he continues, especially important for resource allocators, ICER’s method provides “a very transparent process” for arriving at decisions about where capital should be spent. Weissberg was on the review team that issued ICER’s first evaluation, “CT Colonography for Colorectal Cancer Screening.”

The urgent need for solid information about which medical procedures are most effective for certain patients was highlighted

## IOM Recommendations for Comparative Effectiveness Research

In “Knowing What Works in Health Care: A Roadmap for the Nation,” published in January, the Institute of Medicine made “pretty straightforward” recommendations regarding comparative research, says J. Michael McGinnis, M.D., senior scholar at the IOM. The recommendations “basically focus on the need for greater consistency and capacity” in systematic reviews of evidence, he says.

The IOM’s first recommendation is that Congress direct the Secretary of the U. S. Department of Health and Human Services to designate a single entity to oversee the nation’s work in comparative effectiveness. McGinnis explains that although the report does not identify this single entity, the Secretary might designate the Agency for Healthcare Research and Quality or the National Institutes of Health, another existing organization, or a new entity. The Secretary should appoint an advisory board to oversee the program, including “representation of diverse public and pri-

ivate sector expertise and interests,” the IOM report says.

The responsible entity should appoint a Priority Setting Advisory Committee, also representative of a “broad mix of expertise and interests,” which would “identify high priority topics for systematic reviews of clinical effectiveness,” the report says. McGinnis adds, “An important reason to set priorities is not only to make sure that the most pressing needs are engaged early, but to better coordinate scarce resources” so that several assessment organizations are not working on the same topic.

The IOM’s third recommendation is that “the Program should develop evidence-based methodologic standards for systematic reviews, including a common language for characterizing the strength of evidence.” McGinnis explains, “Inherent in this component is the suggestion that there be a mandate to develop better methods” for evaluating evidence. “There is a little bit of sub-

jectivity when you are weighing the relative merits of one study over another. No Holy Grail has been identified to weigh the relative merits of a study.” This recommendation seems to call for “a formal standard-setting body,” he says.

In the IOM’s final recommendation, the report says, “Groups developing clinical guidelines or recommendations should use the Program’s standards, document their adherence to the standards, and make this documentation publicly available.” The IOM says “a balance of competing interests and diverse stakeholders” should develop the guidelines and that all stakeholders should follow those guidelines.

Says McGinnis, “This is fundamentally a communication challenge—trying to ensure that when guidelines are developed by expert groups, such as thoracic surgeons and cardiologists, that they are using similar standards and are making the best use of the best insights for communication strategies.”

in a February report in the *Journal of the American Medical Association* on the dramatic rise in expenditures for back care, Weissberg points out. The report found that spending on spine treatments in the United States rose 65 percent from 1997 to nearly \$86 billion in 2005, after adjusting for inflation. Despite those rising expenditures, the proportion of people with impaired function due to spine problems increased during that period, even after controlling for an aging population, the *JAMA* report showed.

ICER's work is predicated on the need to include an economic component in any assessment of comparative effectiveness. That component is not often (or not at all) included in assessments produced or sponsored by other national organizations, such as the National Institutes of Health, the Department of Veterans Affairs, the Centers for Medicare and Medicaid Services, and AHRQ.

In testimony before the Subcommittee on Health, Committee on Ways and Means, U.S. House of Representatives in June 2007, Peter R. Orszag, director of the CBO, said, "A reason for the limited availability of information on comparative effectiveness is that public sector health insurance programs—which collectively account for about half of all health care spending—have not sought to make extensive use of it. In particular, 'the Medicare program has not taken costs into account in determining what services are covered and has made only limited use of comparative effectiveness data.... Increasing the amount of credible and objective research that was available could facilitate moving Medicare toward what former program administrator Mark McClellan has called a 'fee-for-value' system rather than a fee-for-service one."

Outlining the federal government's history of involvement in comparative effectiveness research, Orszag mentioned work by "the short-lived National Center for Health Care Technology," which was established in 1978 with a "broad mandate to conduct and promote research on health care technology," and which "ceased operations in 1981, reflecting both changes in priorities for the new administration and the Congress as well as opposition from some provider groups." Similarly, the Office of Technology Assessment, which "studied a variety of health care topics, including the costs and benefits of screening tests for several diseases. . . was eliminated in 1995." Some AHRQ-sponsored studies have analyzed the cost-effectiveness of different treatments as well as their clinical effectiveness, Orszag pointed out.

In the December 2007 CBO report on comparative effectiveness, the agency discussed "the arguments both for and against having federally sponsored research on comparative effectiveness consider cost-effectiveness as well as clinical effectiveness." The agency concluded, "On balance, research that included an analysis of cost-effectiveness would probably have a larger

effect on medical practice than research that analyzed only the comparative clinical effectiveness of different treatments—primarily because the results would sometimes highlight that benefits were small relative to the incremental costs."

Although the IOM recommendations do not include adding an economic component to evaluation of comparative effectiveness, McGinnis says, "Somebody will do it. There will clearly be a fairly substantial cottage industry that will grow up around the scientific determinations to make economic assessments. The work of ICER is quite important because people are going to be doing economic assessments, and it's important there be a dedicated effort to make sure those assessments are using the best methodology. ICER is dedicated to refining the method of cost-effectiveness studies."

### **ICER Gets to Work**

ICER's first two assessments focused on Intensity-Modulated Radiation Therapy (IMRT) for localized prostate cancer and CT Colonography (CTC), or computed tomography colonography, often called "virtual" colonoscopy, a minimally invasive method to screen for colorectal polyps and cancer. In traditional optical colonoscopy, the patient is sedated and a colonoscope, a flexible tube with a video camera on the tip, is inserted into the large intestine. No sedation is required for CTC, in which the patient receives a CT scan while a small tube is used to inflate the bowel. Preparation for both screenings require the patient to empty the bowel with laxatives—one reason many individuals say they avoid the test.

ICER chose these procedures for its first evaluations because while they are drawing a lot of interest, their effectiveness—compared with other available treatments for the same conditions—had not been determined. "You have to prioritize and focus on the things that are most important," says Robert McDonough, M.D., head of Clinical Policy Research and Development at Aetna, who was on the CTC assessment team.

Current screening guidelines recommend colonoscopy for individuals 50 and over as one of the best ways to screen for polyps and cancer. In March, the American Cancer Society included CTC on its list of recommended screening options, but the U.S. Preventive Services Task Force recommendation does not currently include CTC. CTC "is a technology that could potentially be applied to everyone age 50 and older to receive screening for colorectal cancer. You have this new technology and you have an existing technology that's considered the gold standard—optical colonoscopy—so you do need to make a comparative review," McDonough says. According to the ICER report, interest in CTC "is colored by uncertainty over the evidence on the accuracy of CTC, and by questions about the potential impact broad adoption of CTC would have on systems of care and on health care costs."

## AHIP said that work in comparative effectiveness should set “a national research agenda that addresses known gaps in evidence and makes communication regarding ongoing research studies a priority.”

Following current screening guidelines for cancer prevention, most insurers cover traditional colonoscopy. Several insurers cover CTC when, for some reason, a patient cannot undergo a colonoscopy.

ICER’s approach is to perform a systematic analysis of the medical literature on a procedure, using the Evidence-Based Medicine matrix. It also performs “decision analysis modeling” that will give decision makers—health plans, clinicians, and patients—more information on both the clinical effectiveness and the comparative value of the procedure. ICER assigns separate ratings for effectiveness and value but also brings them together in an “integrated evidence rating.” These ratings look somewhat like bond ratings, with combinations of letter grades for effectiveness and value such as Aa, Bb, and so forth.

The bond rating format “helps clarify and structure the discussion of comparative effectiveness by putting it into a simplified model that focuses on two different factors: the comparative value and the comparative clinical effectiveness, so that one of the things you would consider is whether the evidence seems to indicate that a technology is superior,” McDonough says.

In its report on CTC that was completed in January 2008, ICER summarized its systematic review of the medical literature, which, based on the evidence available at that time, concluded that CTC appears to be a safe and effective procedure to screen for colorectal polyps and cancer. One subsequent publication has suggested that there may be some cancers that would be difficult for CTC to identify. But when ICER’s report was published, the review “revealed evidence that newer machines and improved radiologist training have led to robust findings of comparable accuracy to colonoscopy.” Although a small majority of patients would prefer the noninvasive procedure, “it is unclear whether the preference...would result in a larger number of unscreened individuals in a population becoming screened,” the report says.

ICER performed economic modeling to arrive at a comparative value rating for a strategy of screening with CTC every five years—versus no screening—and referring for colonoscopy all lesions greater than a specific size. “Economic modeling shows that, compared to no screening, CTC screening would save lives, and the cost for each additional year of life saved would be approximately \$1,500,” the report says. Traditionally, policymakers have viewed new interventions that produce gains of a year of life at a cost of less than \$50,000 to \$100,000 as a good or reasonable value.


In a head-to-head evaluation, ICER also compared a strategy

of CTC every five years versus colonoscopy every 10 years. Because CTC would have to be performed twice as often, the report concludes that “Compared directly to colonoscopy, if CTC were reimbursed at the same price, CTC would save an additional year of life at a cost of approximately \$600,000. If reimbursed at half the price of optical colonoscopy, CTC saves an additional year of life at a cost of approximately \$100,000.”

Based on this analysis, ICER’s Integrated Evidence Rating was as follows: The Comparative Clinical Effectiveness of CT colonography for colorectal cancer screening versus no screening was rated as: A – Superior; its comparative value was rated as: a—High. The Integrated Evidence Rating versus no screening was therefore: Aa. But when CTC every five years was compared head to head with traditional colonoscopy every 10 years, the comparative clinical effectiveness received a “C” rating (comparable), and only if the reimbursed price was approximately half the price of colonoscopy did CTC receive a comparative value rating of “b” for reasonable value.

“There’s an assumption that virtual colonoscopy will improve compliance,” McDonough says, “but we’re not able to reach any conclusions about the effect of virtual colonoscopy on compliance because when we look at what affects patient compliance with standard colonoscopy, the most troublesome aspect of the procedure is not the colonoscopy itself but the cathartics (preparation).” He adds, “Within 12 months or so, they will have developed a noncathartic regimen, which will allow you to undergo virtual colonoscopy without bowel cleaning, so even without direct evidence we could say that virtual colonoscopy will likely result in greater compliance.” A reassessment of the effect of virtual colonoscopy on compliance may be needed when the noncathartic regimen is developed, he says.

Insurers including Kaiser have begun posting ICER’s CT Colonography report on their internal Web sites. Washington State officials used the assessment in determining colonoscopy coverage policy decisions for state programs, and the report is posted on the state Health Care Authority Web site.

ICER plans to move on to assessments of proton beam therapy and brachytherapy for prostate cancer to determine the clinical effectiveness and value of both procedures. It is work that is urgently needed. Says the IOM’s McGinnis, “In the spectrum of health policy challenges, whether you talk about coverage or quality or cost or prevention or a range of issues, the central challenge to the delivery system is improving the efficiency of its investment—getting better value for our investment.” 

---

**Louise Kertes** is a contributing editor for *AHIP Coverage*.