

Issue Brief

NOVEMBER 2017

KEY POINTS

- Comparative effectiveness research is defined by the Institute of Medicine as “the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care.”
- In the United States, comparative effectiveness research typically does not take into consideration the cost of a given treatment; however, there are international case studies that include cost-effectiveness analysis to help illustrate how implementation can help moderate the growth in health care spending.
- Actuarial models may be able to assist in changing the delivery system to a system that allows consumers, providers, and payers to evaluate the efficiency and quality of care through the use of comparative effectiveness.



AMERICAN ACADEMY of ACTUARIES

Objective. Independent. Effective.™

1850 M Street NW, Suite 300
Washington, DC 20036

202-223-8196 | www.actuary.org

Craig Hanna, Director of Public Policy
Cori E. Uccello, Senior Health Fellow

© 2017 American Academy of Actuaries. All rights reserved.

Comparative Effectiveness Research

Executive Summary

Health care costs continue to grow at a pace that is considered by most standards to be unsustainable, especially in the United States. Many proposed solutions shift costs between those who fund health care and those who receive it. However, these proposals do not affect the rising cost of health care, only who pays for it.

Research on ways to change the underlying cost of health care is ongoing; the objective would be to apply this research to the delivery and coverage of health care. One area is comparative effectiveness research (CER), which the Institute of Medicine (IOM)¹ defined as “the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care.”

CER focuses on which treatment works best, for whom, and under what circumstances, without consideration of the cost of treatment. While focused on improving outcomes, CER also has the potential to be used to change the underlying cost of health care. In many circumstances, the ability to choose the most appropriate treatment may reduce health care costs in addition to improving health outcomes.

In light of the growth of spending for U.S. health care, which continues to grow at a faster rate than the gross domestic product (GDP), the amount of research into the value of health care interventions is also increasing. For example, entities such as Kaiser Permanente’s Division of Research and HealthCore (a collaboration between Eli Lilly and Anthem) have formed to study the cost effectiveness of prescription drugs.²

¹ In March 2016, the IOM was renamed the Health and Medicine Division of the National Academies of Sciences, Engineering, and Medicine to indicate an increased role in broader health care issues.

² Recent publications include: “[Oral Steroids for Acute Sciatica Produce Limited Improvement in Function and Pain](#)” from the Kaiser group and “[Certain Patients with Type 2 Diabetes Less Likely to Suffer Heart Failure While Taking New Class of Antidiabetic Drug](#)” from members of the HealthCore group.

The American Academy of Actuaries' Comparative Effectiveness Work Group of the Health Care Delivery Committee developed this issue brief to provide an actuarial overview of comparative effectiveness research, including measuring its impact and its potential to reduce health care costs. This issue brief includes how CER is conducted, examples of U.S. and international organizations that employ CER, and actuarial and public policy considerations.

How Comparative Effectiveness Research Is Conducted

In addition to the IOM's definition of comparative effectiveness research, it also states that "[t]he purpose of CER is to assist consumers, policymakers, clinicians, and purchasers to make informed decisions that will improve health care at both the individual and population levels."³ CER is a broad term, and several methods may be used to conduct this research. The most common methods include:

- *Systematic Review.* A systematic review is a critical assessment and evaluation of all research studies that address a particular clinical issue. The researchers use an organized method of locating, assembling, and evaluating a body of literature on a particular topic using a set of specific criteria. A systematic review typically includes a description of the findings of a collection of research studies. The systematic review may also include a quantitative pooling of data, called a meta-analysis.
- *Randomized Controlled Trial.* Under the randomized controlled trial method, participants are randomly assigned to two or more groups that differ only on the basis of exposure to the study variable addressing the clinical question (namely, the medications, procedures, or diagnostic tools being compared). The groups are followed for predetermined outcomes of interest to address the question at hand, and the results of the groups, as measured by quantitative metrics, are compared by statistical analyses.
- *Observational Study.* In an observational study, participants are not randomized or otherwise pre-assigned to a treatment. The choice of treatments is made by patients and their physicians. Observational studies can be prospective or retrospective. Prospective observational studies are observational studies in which the outcomes are studied after the creation of a study protocol and analysis plan. The intervention may include surgery, changes in exercise or diet, use of medical devices, use of prescriptions, etc. This contrasts with a retrospective observational study that uses existing data sources, such as claims data or medical records, in which both the intervention and outcomes have already occurred.

³ See, for example, "[Defining Comparative Effectiveness Research: The Importance of Getting It Right](#)," *Medical Care*; June 2010.

Members of the Comparative Effectiveness Work Group include: Susan Pantely, MAAA, FSA, chairperson; Jeff Adams, MAAA, ASA; Colleen Driscoll, MAAA, FSA, FCA, EA; Audrey Halvorson, MAAA, FSA; Malgorzata Jankowiak-Roslanowska, MAAA, ASA.

Factors Influencing Choice of Method

Each of the methods has strengths and limitations. Randomized controlled trials are often considered the gold standard for clinical research and are ideal for research that requires a high degree of certainty. However, randomized controlled trials can be expensive, labor-intensive, and time-consuming. Additionally, randomized controlled trials are particularly difficult to conduct when studying rare diseases due to the limited number of patients that can participate.

Observational studies are typically faster and more cost-efficient. For studies of rare diseases, observational studies can provide the volume of data needed for statistical significance that randomized controlled trials often cannot.

Observational studies are also a useful method when trials have not been or cannot be performed. For example, there are instances when it would be ethically unacceptable to deny access to an intervention or to deliberately expose patients to less-effective treatments. However, observational studies rely on claim data analyses and other historical sources where data quality issues and other data limitations exist.

Observational studies use an abundance of data that is often essential when treatment effects differ across types of patients and when analyses of subgroups are needed to understand which patients are most likely to benefit. Regional differences in results due to preferences for various treatment options that occur across locations may also be captured in observational studies.

Actuarial Considerations for Comparative Effectiveness Research

Actuarial practice has always involved retrospective observational studies using claims data, most often to understand drivers of costs and trends. These same actuarial techniques can be used for CER. Historically, actuaries have relied on health care claims data for many purposes, such as estimating future premium

rates and studying the impact of cost-sharing on utilization of services. Actuaries generally pay close attention to data issues, as data quality will affect the accuracy of their conclusions. Health care data almost always has known or unknown limitations, and medical claim and demographic data may be influenced by environmental changes, technology changes, and other factors.

Additionally, claims data typically does not include other information that may be useful such as patient vital statistics, lab results, or outcome information. Historically, this type of information has been unavailable; however, this data is becoming more attainable with the development of robust computer capabilities and increased interest in studying outcomes through observational studies. As medical claims data begins to incorporate clinical and medical-chart data, actuarial techniques can be adapted to include these variables into observational studies. Privacy concerns, however, on data at this level of granularity may hinder wide adoption of compiling this information. In addition, members switching to other health plans can affect claims data and limit the ability to do longitudinal studies.

Retrospective observational studies must be designed to avoid selection bias. For example, if physicians prescribe one treatment over another based on the severity of the condition, the data will be biased. Two tools actuaries use to address selection bias are risk adjustment and propensity score:

- *Risk Adjustment.* Risk adjustment is an actuarial tool that identifies a risk score for a patient based on the conditions identified via claims or medical records. Risk adjustment can be used to calibrate payments to health plans or other stakeholders based on the relative health of the covered population. It can also be used to identify similar types of patients for comparative purposes.

Prospective risk adjusters use an individual's medical claims data to predict his or her future costs. Concurrent risk adjustment is not a true predictive model but rather uses medical claims data to explain an individual's current costs. The study design will have to decide whether prospective or concurrent risk adjustment models are more appropriate depending on the situation.

- *Propensity Score.* A propensity score is the conditional probability of receiving treatment given several variables that may be predictive, including the values of all treatment potential biases. Patients in a treatment group are matched to control group patients on the basis of their propensity score. Differences in outcomes are estimated between balanced patient groups. Regional differences in treatment options may be controlled through reliance on a patient group's propensity score.

Comparative Effectiveness Research Can Be Used For Medical and Prescription Drug Treatments

Comparative effectiveness research is often performed to assess quality of care and efficiency of prescription drug treatment, non-drug treatments, and all medical treatments combined, including drug and non-drug treatments.⁴

Arguments can be made for and against including both medical and prescription drug treatments in a comparative effectiveness analysis. An analysis that only includes prescription drug treatments may be easier to undertake because the treatments being analyzed are similar in nature (i.e., an individual taking prescription drugs for a particular medical condition). The methodology used and expertise required in reviewing each of these drugs would be similar. If drug and non-drug treatments are to be

studied in the same comparative effectiveness analysis, the analysis becomes more complicated, and the expertise required to perform the analysis becomes more comprehensive. The use of a prescription drug-only analysis may be appropriate if non-drug treatments are not available or if prior studies have shown that drugs are the best option. There also may be situations where not enough claims experience exists to determine the effectiveness of certain treatments.

Although more complicated, analyses using both prescription drug and non-drug treatments may be more useful if both drug and non-drug treatments are available. For example, some studies have indicated that a change in diet and exercise habits is a better treatment for Type 2 diabetes than prescription drugs.⁵ If this were true, then a prescription drug-only analysis would fail to produce the most effective treatment for Type 2 diabetes.

Comparative Effectiveness Research Is Not the Only Approach

There are other approaches that incorporate the cost and value of medical treatments, including cost-benefit (assigning a monetary value to an outcome) and cost-effectiveness (comparing the relative benefits of two or more options). Cost-benefit analysis is typically measured as “cost per life-year saved” or “cost per quality-adjusted life-year saved.”

Several related types of research are often confused with CER. For example, medical efficacy studies are used to determine the degree to which an intervention accomplishes the desired or projected outcomes. Because medical efficacy studies do not compare different treatments to determine which works best, they are not considered CER studies.

⁴ For example, New Zealand has described its methodology for analyzing the effectiveness of prescription drugs. See [Prescription for Pharmacoeconomic Analysis: Methods for Cost-Utility Analysis](#); Pharmaceutical Management Agency; 2012. Other organizations such as Cochrane and E-Science Central recommend and develop studies, including both drug and non-drug treatments.

⁵ See, for example, [“The Prevention and Control the Type-2 Diabetes by Changing Lifestyle and Dietary Pattern”](#); *Journal of Education and Health Promotion*; 2014.

Comparative Effectiveness Research in the United States

While a number of organizations perform comparative effectiveness research, we highlight two particular examples in this issue brief: the Patient-Centered Outcomes Research Institute (PCORI) and The Dartmouth Institute. These are just two organizations that offer a few case study examples to better explain CER.

Patient-Centered Outcomes Research Institute

PCORI was authorized by the Affordable Care Act (ACA) and established in 2010. Its purpose is to fund patient-centered CER, which means that patients are included in choosing the research questions important to them. PCORI focuses on questions that patients have related to their own particular situation and results that are meaningful to them.⁶

The research PCORI supports is not typical randomized controlled clinical trials. PCORI is focused on patient engagement, enrolling many people, collecting data, and measuring outcomes. It uses observational studies, reviews practical application of clinical trial results, and reviews already-published research (systematic review).

The Dartmouth Institute

The Dartmouth Institute (TDI) was founded in 1988 in an effort to research efficiency and effectiveness in health care delivery. Through the Dartmouth Institute CER program and the Dartmouth Atlas Project, TDI analyzes various treatment protocols and effectiveness by region of the country in an attempt to determine those treatments or areas that lead to more effective health care. TDI CER evaluates treatments for diseases and issues; the Dartmouth Atlas Project analyzes regional differences in treatment.

According to TDI, its researchers work to address the following questions:⁷

- “How do medical technologies and innovations in health care delivery improve clinical care and outcomes for patients?”
- How are innovative medical technologies and delivery practices best used for prevention, diagnosis, and treatment of illness?
- How can health information technology be used to provide individually tailored decision support for patients and their health care providers?
- What methodological approaches are best for evaluating what works for disease prevention?”

Application of Research to Current Practice

Comparative effectiveness research often suggests potential actions that can be taken to encourage appropriate use of certain treatments. The findings may also provide opportunities for further research, sharing of information with patients and providers for decision-making, and additional education of health plan members of certain benefits available to them.

The following case studies are examples of PCORI and TDI funded research results:

- *Osteomyelitis in children.* PCORI funded a study researching post-discharge antibiotic therapy administered via the peripherally inserted central catheter versus treatment using an oral antibiotic for acute osteomyelitis in children. The study found that the oral therapy group had fewer complications, and there was a lower risk of complications requiring an emergency room visit or hospitalization. The study suggests that clinicians may wish to consider using

⁶ For more information, see <https://www.pcori.org/>.

⁷ For more information, see <http://tdi.dartmouth.edu/>.

- oral antibiotics for the treatment of otherwise healthy children with acute osteomyelitis rather than prolonged intravenous antibiotics after hospital discharge.⁸
- *Warfarin study.* Warfarin is used to prevent major adverse cardiovascular events in patients with atrial fibrillation after stroke. Clinical trials have shown value, but PCORI sponsored an observational study of the practical use of warfarin in Medicare patients with atrial fibrillation who were admitted to the hospital for an ischemic stroke. The results of this study showed improved outcomes for these patients using warfarin as compared to those not prescribed an oral anticoagulant at discharge. The additional value of this study is that it looks at the actual practice of the use of warfarin through an observational study, versus a controlled clinical study.⁹
 - *Supplemental breast cancer screening.* Some states require women with mammographically dense breast tissue be notified that they may benefit from further screening beyond mammography, including breast ultrasonography and 3-D mammography. Researchers at TDI engaged in modeling with three Cancer Intervention and Surveillance Modeling Network (CISNET) simulation modeling groups. TDI found that ultrasonography would increase costs with little benefit in terms of health outcomes; however, they recommend further study for supplemental 3-D mammography because there is evidence of potential health benefits without a significant increase in cost.
 - *Spine treatment outcomes calculator.* Two of the most frequently performed surgeries are related to lower back pain; however, the decision to perform these surgeries vary broadly by geographical location.

Using evidence from clinical trials and in collaboration with *Consumer Reports*, TDI's Multidisciplinary Clinical Research Center in Musculoskeletal Disease developed a web-based calculator to help people compare potential outcomes (benefits and harms) associated with surgery for three specific problems—intervertebral disc herniation, spinal stenosis, and degenerative spondylolisthesis.

International Comparative Effectiveness Research

While the United States has only recently started focusing on CER, other countries have been using CER as well as cost-effectiveness research to make coverage and treatment decisions for many years, including the United Kingdom and France. The following section examines some of the CER research being done in these other countries, including case studies.

United Kingdom

The U.K. started using CER in 1999, when it formed the National Institute for Health and Clinical Excellence (NICE). The organization does not conduct medical research but does commission research in the form of evidence synthesis and economic modeling by other academic research centers and medical centers. When NICE started, the research only focused on clinical effectiveness. But in 2006, it changed its function to add cost-effectiveness.

Research is submitted to NICE in the form of medical case studies. These are assessed for savings, quality, effectiveness, and implementability. Each case study must show how the new drug or procedure costs less with the same result or medical results are improved at the same costs. A study must also explain how to implement the new technology, offer a timeline, and provide information on potential barriers to widespread implementation.

⁸ "Comparative Effectiveness of Intravenous vs. Oral Antibiotics for Postdischarge Treatment of Acute Osteomyelitis in Children"; *JAMA Pediatrics*; February 2015.

⁹ "Real world effectiveness of warfarin among ischemic stroke patients with atrial fibrillation: observational analysis from Patient-Centered Research into Outcomes Stroke Patients Prefer and Effectiveness Research (PROSPER) study"; *British Medical Journal*; July 31, 2015.

NICE can respond to the research studies on a new technology, treatment, or drug in four ways:

- *Recommended.* In this instance, the new treatment/drug is approved to be in line with marketing, and in line with noted clinical practice. This then moves to the implementation phase.
- *Optimized.* In this case, the new treatment/drug is approved but only to be provided to a smaller subset of the patient population. This is usually due to cost.
- *Only in research.* NICE can put the new technology, treatment, or drug into a clinical trial in order to gather more data.
- *Not recommended.* This is either due to lack of evidence of effectiveness or because the technology is not cost-effective.

Once a new drug or technology is recommended, it moves to the implementation phase. NICE publishes guidance to the National Health Service, which then takes the guidance and makes the new drug or technology available and acts as the enforcer for the new guidance.

Case Studies

- *Thyroid cancer test kit.* The case study explained that current fine-needle aspiration mutation testing methods were sometimes inconclusive—they would fail to indicate whether the thyroid was cancerous or not. When a failure occurred, surgery was required to get a larger sample. At that time, if the results were cancerous, a second surgery was needed to remove the entire thyroid. The new testing kit was 99 percent accurate with the fine-needle aspiration. Therefore, the need for the surgery to gather the larger sample was no longer needed. The trial ran for 18 months and tested 57

samples, which resulted in the avoidance of 11 surgeries. The study provided a cost comparison of the 57 tests versus the 11 avoided surgeries and showed a savings of approximately \$17,500. The study provided information on increased quality and increased patient satisfaction because of quicker test results and the avoidance of surgery.

The implementation was estimated at four to 12 months. This timeline was relatively fast because the new commercial test kits can be made widely available quickly. Barriers noted were the need for local expertise and equipment for running the test.¹⁰

- *Diuretics for heart failure.* This case study was first published as a “proposed example.” After the research team added more evidence, this became the official guidance a year later. The research involved providing heart failure patients with intravenous diuretics at home rather than admitting them to the hospital. Results from all 10 pilot sites were positive. As the average hospital stay in these situations is 13 days, the study showed savings by avoiding 869 bed-days over the two-year testing window. The quality also was better, as patients said they were happier staying at home with nurses who provided home care.

The implementation was estimated at six to 12 months, and the challenge was to get buy-in from cardiologists and heart-failure nurses early in the process because they would have to assess whether candidates were suitable for home therapy.

¹⁰ [BRAJV600E mutation testing for thyroid cancer: avoiding unnecessary surgery](#); Portsmouth Acute Hospitals NHS Trust; December 2015.

France

The French National Authority for Health (Haute Autorité de Santé, or HAS) was established in 2005 as an independent public body whose mission includes contributing to the regulation of the health care system by providing information on health quality and efficiency to decision makers. In particular, this mission includes health economics assessments and opinions on the most efficient strategies for health care and prescribing.

HAS carries out technology assessments that analyze a new technology's intrinsic benefit and its effectiveness compared with that of existing technologies. A single technology assessment is required before a new drug, device, or medical procedure can be added to the benefit list for sickness funds. A new treatment may not be covered unless it provides either improved benefits or lower costs. HAS recommendations are advisory; however, its findings are typically accepted by the Ministry of Health or the union of sickness funds.¹¹

Focusing on economic evaluations, HAS developed a set of guidelines to formalize reference case analyses in economic evaluations.

Case Studies

- *Assessment of an edge-to-edge mitral valve repair clip and its implantation.* Mitral insufficiency is the second-most common valve disease in Europe. The study compares the MitralClip device to conventional valve repair or replacement surgery, and included data from the manufacturer, a literature search, clinical guidelines, and technological assessments. HAS concluded there were no alternatives for patients with severe degenerative mitral insufficiency, which is symptomatic despite optimal medical treatment. Improvement for these patients was substantial in relation to lack

of alternatives. Thus, HAS recommended the use of the MitralClip in these limited circumstances. However, HAS recommends the MitralClip implantation be supervised due to the difficulty of the technique for implanting an edge-to-edge mitral valve repair clip, the resulting learning curve, and the strict patient selection. Finally, HAS also recommended that a registry be established to track all patients undergoing this procedure.¹²

- *Fluenz Tetra.* Fluenz Tetra is a live attenuated, nasal vaccine with marketing authorization in the prevention of influenza in children over age 2. The study further discusses the clinical aspects of the vaccine. The actual benefit of Fluenz Tetra is substantial, and HAS recommends inclusion on the list of reimbursable products for supply by pharmacists and for hospital use.¹³

Conclusion

Examples from other countries show the value of implementing CER to help change the cost of care delivery, which could moderate the growth in health care spending. Although in some cases these studies include cost-effectiveness analysis, which is not yet widely used in the United States, these country-specific studies also offer implementation advice.

CER is relatively new in the United States; in 2008, a committee of the Institute of Medicine recommended a national program of CER.¹⁴ Its report addressed the problems of multiple, conflicting practice guidelines of widely varying quality and application by region. Widespread adoption will likely be driven by early successes of CER recommendations.

11 [National Authority for Health: France](#); The Commonwealth Fund; July 2009.

12 *Evaluation d'un clip de réparation mitrale bord à bord et de son acte d'implantation: Service d'évaluation des dispositifs*; Haute Autorité de Santé; 2015.

13 See https://www.has-sante.fr/portail/jcms/c_2035692/en/fluenz-tetra-tetravalent-vaccine-against-seasonal-influenza-in-children

14 "Knowing What Works in Health Care: A Roadmap for the Nation"; Institute of Medicine, January 2008.

To date, prompt adoption of CER recommendations has been mixed. Some studies have achieved prompt response, such as hormone replacement therapy to prevent heart disease in women where clinical trial results showed increased rates of heart attacks and other adverse results. Another example where guidelines were quickly adopted is autologous bone marrow transplants for women with breast cancer. Studies showed the autologous bone marrow transplants had no better outcomes than conventional chemotherapy and had much higher risk of serious side effects, and, therefore, these bone marrow transplants were limited.¹⁵

However, some studies have not achieved wide adoption. Research on prostate-specific antigen (PSA) screenings, for example, show that screenings offer little benefit; however, this test is still widely used. Another example the use of is thiazide diuretics for patients with hypertension, which has been shown to be superior in preventing cardiovascular disease events and less expensive.

While the use of CER can drive change in the delivery and underlying cost of health care, the funding of health care is vastly different among countries. The adoption of a CER mindset, which embraces use of widely disseminated and constantly updated results from treatment alternatives, serves to improve patient outcomes and slow the increase in cost of providing the treatment. It may be more difficult to implement CER into health care delivery or insurance coverage in the United States in the commercial insurance world (individual under age 65 and employer coverage) due to the level of regulation and the free-market environment.

What is known is that CER can identify more effective treatment for certain conditions or certain patients, and should be considered as one part of a solution for changing the underlying cost of health care. There are opportunities with government programs to pilot the use of CER in health insurance coverage. Marketing CER successes will aid in its widespread adoption. Providers may also wish to embrace CER in their own protocols and medical policy programs. Actuarial models may be able to assist in changing the delivery system to a system that allows consumers, providers, and payers to evaluate the efficiency and quality of care through the use of comparative effectiveness. This could be done by various mechanisms such as changes in provider payments or changes in benefit design to encourage the use of the more effective treatment.

¹⁵ [“The Controversy Over High-Dose Chemotherapy With Autologous Bone Marrow Transplant for Breast Cancer”](#); *Health Affairs*; September 2001.

The American Academy of Actuaries is a 19,000-member professional association whose mission is to serve the public and the U.S. actuarial profession. For more than 50 years, the Academy has assisted public policymakers on all levels by providing leadership, objective expertise, and actuarial advice on risk and financial security issues. The Academy also sets qualification, practice, and professionalism standards for actuaries in the United States.