The Facts about Comparative Effectiveness Research
How Studying Which Treatments Work Can Improve Care and Reduce Costs
The Facts about Comparative Effectiveness Research
How Studying Which Treatments Work Can Improve Care and Reduce Costs

Jeffrey Bernstein
U.S. Public Interest Research Group

WISPIRG Foundation

July 2009
Acknowledgments

The author thanks Brent James, M.D., M. Stat., Executive Director, Institute for Healthcare Delivery Research, Intermountain Healthcare, and Professor Douglas L. Wood, M.D., Mayo Clinic, for making themselves available for inter-views for this report.

Thanks to Tony Dutzik of Frontier Group, Mike Russo of CALPIRG, Laura Etherton of OSPIRG, and Steve Blackledge and Larry McNeely of U.S. PIRG for their editorial assistance.

The author bears responsibility for any factual errors. The recommendations are those of WISPIRG Foundation. The views expressed in this report are those of the author and do not necessarily reflect the views of our funders or those who provided review.

© 2009 WISPIRG Foundation

With public debate around important issues often dominated by special interests pursuing their own narrow agendas, WISPIRG Foundation offers an independent voice that works on behalf of the public interest. WISPIRG Foundation, a 501(c)(3) organization, works to protect consumers and promote good government. We investigate problems, craft solutions, educate the public, and offer Americans meaningful opportunities for civic participation.

For more information about WISPIRG Foundation or for additional copies of this report, please visit: www.wispirg.org.

Cover photo: Fotek, licensed by iStockphoto.com
Layout: Harriet Eckstein Graphic Design
# Table of Contents

Executive Summary | 1
--- | ---
Introduction | 3
Lack of Evidence-Based Care in Medicine Today | 5
Examples of Comparative Effectiveness Research | 7
- Turpentine and Bezoar Stones | 7
- Emphysema Treatment | 7
- Coronary Artery Disease | 8
- Antibiotics for Surgical Wounds | 8
- High Blood Pressure | 9
Current Uses of Comparative Effectiveness Research | 10
- Mayo Clinic | 10
- Intermountain Healthcare | 11
The Need for Further Comparative Effectiveness Research | 12
Federal Support for Comparative Effectiveness Research | 14
Conclusion | 16
Notes | 17
As Congressional and public debate over health care reform grows more intense, comparative effectiveness research (CER) has emerged as an unlikely flashpoint of controversy. Opponents’ claims that CER results in the rationing of health care or a government takeover are belied by the true nature of such research: it is simply fundamental scientific research of medical treatments aimed at determining the most effective ways to treat sickness and injury. It is the basis of all advancements in the field of medical science and has been used throughout history to improve medical treatment. The results of such research are used to create treatment guidelines, which are then incorporated by physicians in determining the best course of care for each individual patient.

The tremendous need for this research is made clear by studies showing that only a minority of medical treatments currently being used are supported by valid research. As medical innovation accelerates, with new techniques, drugs, and devices entering the market daily, the need to gauge the effectiveness of these innovations increases.

Failure to use effective treatments results in worse medical outcomes and higher medical costs, resulting from the need to re-treat patients or to address complications following ineffective treatment. Evidence-based medical protocols, developed from the findings of CER, have been shown to yield cost savings and improvements to patient health.

A recent report released by the Institute of Medicine of the National Academies concluded that “the country needs a robust CER infrastructure” and listed the top 100 priority topics for comparative effectiveness research.1

Opponents claim that CER will mandate doctors’ treatment options. Generating data on effective medical treatment is a separate issue from the use of that data. These studies are a tool that can assist doctors in determining the best treatment for their patients, in conjunction with their own knowledge of the unique needs of each patient. In some cases, the results of CER indicate that there is in fact a best practice that should be followed in nearly all circumstances. In others, research can help establish which kinds of patients are most likely to benefit from which treatment options. Doctors and care providers should
be encouraged to keep current with effectiveness research findings and use them to improve patient care.

Much of this research today is conducted or paid for by the manufacturers of the drugs or medical devices being tested. Such research has been shown to be biased toward a finding that the drug or device is beneficial. Thus it is necessary that CER be funded by neutral parties who do not have an economic interest in the result. The federal government is ideally situated to perform such research and to fund nonprofit institutions to do it.

To improve patient care and reduce the costs of unnecessary and improper treatment, the federal government should expand its funding and support for comparative effectiveness research.
Funding for comparative effectiveness research (CER) is much in the news as one of the key initiatives included in the health care reform plans being formulated in Congress.

Comparative effectiveness research is just what the words suggest: research that compares the relative effectiveness of two or more techniques, drugs, procedures, or medical devices.

The recently passed federal economic recovery package allocated $1.1 billion for CER, the purpose of which is to “conduct, support, or synthesize research that compares the clinical outcomes, effectiveness, and appropriateness of items, services, and procedures that are used to prevent, diagnose, or treat diseases, disorders, and other health conditions.”

The *New England Journal of Medicine* described CER as follows:

In CER, groups of patients are analyzed to compare the effectiveness of alternative medical strategies, with the intent of informing clinical decisions and policies affecting health care....CER offers a way to hasten the discovery of the best approaches to personalization, providing more and better information with which to craft a management strategy for each individual patient.

CER is simply the use of the scientific method—observation, hypothesis, prediction, experiment—within the realm of medicine. Strictly defined, CER involves rigorous experiments called randomized control trials in which a medical procedure, drug, or device is used on one set of people while a second group, known as the control group, is given an alternative treatment, a placebo, or no treatment at all. The patient does not know which group s/he is in, and often the personnel interacting with the patient do not know, either. The results from the two groups of patients are analyzed to see if there is a statistical difference in the outcome. If there is such a difference, it shows that one treatment is more effective than the other. This type of research can also be less formalized, as examples below describe, where changes in treatment methods show immediate and substantial improvement with no adverse consequences.

The results of comparative effectiveness
research studies may be used to produce clinical guidelines for care. These guidelines, known as “evidence-based protocols,” are not absolute mandates—they are treatments or methods that have been shown to be effective in the majority of cases. The guidelines are intended as a starting point from which doctors can determine the proper care for each particular patient. They also provide a benchmark by which physicians, insurers, and other professionals can determine if patients are receiving basic, appropriate care.

In some cases, these protocols are specific, for example where CER has determined the optimal time to use a particular drug or treatment. In other cases, studies show a wider variety of outcomes, meaning that the protocols offer more options as to the appropriate kinds of treatment. Evidence-based protocols developed according to this process have been shown to improve health outcomes, while reducing the costs of care.

In the words of Harvard and Michigan Medical Schools professors Michael Chernow and Mark Fendrick, “CER is an essential tool to determine which intervention should be delivered to which person and in what clinical circumstance.” CER is a tool—one of many—that a physician uses to determine the best care for his/her patient.

Comparative effectiveness research provides evidence as to what the most effective treatments are for medical conditions. How that information is used is up to each doctor, nurse, pharmacist, or clinician. Typically the doctor using CER will follow the indicated treatment unless his/her knowledge of the patient leads the doctor to believe that the particular treatment will not be as effective for that particular patient.
null
The study found that patients were only receiving about one-half of the recommended treatments. The percentages were nearly the same for all types of care: acute care, 53.5%; chronic care, 56.1%; preventative care, 54.9%. While the problem was usually that patients did not receive the recommended care, in more than 11% of the cases patients received too much care, care that “was not recommended and was potentially harmful.”

To cite a specific example, out of every 1,000 Medicare enrollees in Bangor, Maine, 0.2 were treated with lumbar fusion to alleviate back pain while in Idaho Falls, Idaho, the lumbar fusion treatment rate was 4.6 per 1,000, “with no difference in the outcomes.”

People with back pain in Idaho Falls were treated with lumbar fusion twenty times more often than in Bangor, but those people who had lumbar fusions were no better off. Too many people are receiving care without evidence that it is beneficial.

Thus we have a health care system in which many treatments are not proven to be effective, and in which patients often do not receive treatments known to be appropriate. Failure to treat properly results at best in having to treat again and at worst in more serious complications. Overtreatment, the use of unnecessary tests and procedures, adds to costs without any benefit to patient health.
Examples of Comparative Effectiveness Research

Comparative effectiveness research, in one form or another, has been around throughout the history of modern medicine.

**Turpentine and Bezoar Stones**

The 16th century French surgeon Ambroise Parè is known for at least two endeavors in the field of comparative effectiveness research. Prior to Parè, battlefield wounds were treated by pouring boiling oil on the wound. Parè instead used a mixture of egg yolk, oil of roses, and turpentine and found a significantly better result (due, apparently, to the antiseptic properties of turpentine).¹¹

In a second experiment, he examined the supposed properties of bezoar stones as an antidote to poisoning. Bezoar stones are masses of undigested food (seeds, pits, vegetable matter) or hair lodged in the digestive tract, typically obtained from goats and cows. It was believed that when a bezoar stone was placed in water, the water would then counteract the effect of any poison. King Charles IX of France acquired a bezoar from Spain and asked Parè, his physician, if there were any universal antidotes as effective as the bezoar. Parè suggested an experiment to test the bezoar’s effectiveness. He asked if there were any prisoners scheduled for hanging. As it happened, a cook had recently been sentenced to death for stealing silverware. The king made the cook a proposition: he could be hanged as planned or he could be poisoned, at which time he would also be given a bezoar. If the bezoar successfully blocked the poison, he would not be hanged. The cook opted to take his chances with the poison.

He was given poison, immediately followed by the bezoar water. The cook died, proving to Parè that the bezoar was not effective as a universal antidote.¹²

Much of the material in medical journals is comparative effectiveness research. Below are several more modern examples.

**Emphysema Treatment**

Emphysema is a condition where the parts of the lungs called the alveoli—the sites
where the oxygen breathed in is transferred to the bloodstream—and the narrow passages leading to them, become brittle or break. This results in air gathering in the lungs but unable to reach the bloodstream.

Prior research had indicated that surgery to reduce lung volume by removing a portion of the lung could be an effective treatment for severe emphysema. Scientists performed a study where surgery was compared with other medical therapies (including prescription drugs and lifestyle changes) among different types of emphysema patients.\(^13\)

While the study found that “[l]ung-volume-reduction surgery was associated with a greater chance of improvement in exercise capacity, lung function, quality of life, and dyspnea [shortness of breath],”\(^14\) it noted that for certain types of people who were in poor physical condition and whose emphysema was of a particular type (upper lobe), the improvement in their capacity for exercise, the improvement in their symptoms, and the decrease in the likelihood of dying, was significantly greater than with medical therapy. However, the surgery was not found to be better than medical therapy for patients with other types of emphysema who were in better physical condition. Thus CER showed which treatment was better for which type of patient.

In the above example, CER compared surgery to drug treatment. In a study on the treatment of blocked arteries, CER compared the use of medical devices to drug therapy.

Coronary Artery Disease

Coronary artery disease is a condition in which the arteries that bring oxygenated blood to the heart are constricted by the build-up of obstructions (plaque) within the artery. It is typically treated with drugs (ACE inhibitors), by reducing risk factors (lowering LDL cholesterol, weight reduction, increased physical activity), and through lifestyle change (improved diet, exercise, stopping smoking). This constellation of treatments is regarded as “optimal medical therapy.”

In addition, for the past thirty years doctors have prescribed percutaneous coronary intervention (PCI). PCI involves either or both angioplasty or putting a stent into the artery. Angioplasty involves pushing a balloon through the artery to remove the blockage. A stent is a tube which is placed in the artery to hold it open.

A study was conducted to determine whether optimal medical therapy in conjunction with PCI reduces the risk of death and heart attack more than optimal medical therapy alone in patients with stable coronary artery disease.\(^15\) (“Stable” means that the patient experiences chest pain but the pain does not worsen over time, and the person has not had a heart attack.)

This study, in conjunction with other studies on more than 5,000 patients, proved that neither the angioplasty procedure nor the stent reduced the likelihood of death or non-fatal heart attack as compared with optimal medical therapy alone.\(^16\) “PCI has no effect in reducing major cardiovascular events.”

Another area of research involves determining the most effective timing for the administration of medication.

Antibiotics for Surgical Wounds

Previous research had shown that antibiotics are effective in reducing infections from surgical wounds. But it was not known how the timing of the use of the antibiotics would affect results. This study sought to determine the best timing of the use of antibiotics to prevent surgical infections.\(^17\)

Experimenters studied over 2,800
patients who had undergone surgery at a large community hospital. They divided the patients into four groups—those given antibiotics 2 to 24 hours before the surgery, during the 2 hours before the surgery, during the 3 hours after the surgery, and between 3 and 24 hours after the surgery.

The results clearly indicated that antibiotics should be given within two hours prior to the surgery. Compared to the group receiving antibiotics in the 2 hours before the surgery, those people given antibiotics within 3 hours after surgery were more than twice as likely to develop infections from the surgical incision. Among the patients receiving antibiotics 3-24 hours after surgery, infection was more than 5 times as likely, and it was more than 6 times as likely among the group receiving antibiotics more than 2 hours before the operation. Comparative effectiveness research is necessary to make these sorts of distinctions among treatment options.

Effectiveness research studies are also done between drugs which treat the same conditions to determine their relative effectiveness. One important study compared three types of drugs used to treat hypertension (high blood pressure).\(^{18}\)

### High Blood Pressure

About a third of the adult U.S. population has high blood pressure.\(^ {19} \) While lifestyle changes (stopping smoking, losing weight, getting more exercise, improving diet, reducing stress) help lower blood pressure,\(^ {20} \) over 20 million Americans are treated with prescription drugs for high blood pressure.\(^ {21} \) There had been controversy for many years among doctors as to which drugs best treated this condition. An older set of drugs known as diuretics had been preferred, but doctors, recognizing the possible side-effects of diuretics, had been shifting to two other types of drugs, ACE inhibitors and calcium channel blockers. These latter drugs were significantly more expensive than the diuretics.

The results of this study were “striking,” “robust,” and “unambiguous” that the less-expensive diuretics “should be the initial drug of choice for patients with hypertension.”\(^ {22} \)
Results of comparative effectiveness research are improving treatment in doctors’ offices and hospitals across the country. For example, the Mayo Clinic has used such research to improve treatment in the use of warfarin.

Mayo Clinic

Warfarin is a blood-thinning medication, taken to reduce the likelihood of blood clots which can break loose and block blood flow. It is the most commonly prescribed anticoagulant drug in North America. Despite its effectiveness, treatment with warfarin has several shortcomings. Many commonly used medications and foods interact negatively with warfarin. Also, too much warfarin may result in internal bleeding while too little will not prevent the dangers of blood clots. Thus it is essential that the dosage be correct for the patient.

Traditionally, doctors prescribed warfarin according to their own professional judgment. Mayo found that physicians were not consistent in their decisions about dosage. The Clinic sought to improve its warfarin therapy and thus experimented with new methods of warfarin delivery. They considered the known risk factors: patient’s age, coexisting diseases, and other drugs the patient was taking. This information along with blood tests produced a “standardized protocol or algorithm” which was used to determine the proper warfarin dosage. Doctors compared this new method against their standard approach in small experimental cycles and found that their new method was far more effective at reducing adverse events than what they had been doing previously.

As a result, whereas in 2007 Mayo had 412 instances where patients received too much warfarin (resulting in an average of three extra hospital days each), that number was reduced to 158 in 2008 and there have been none in 2009 as of this writing. Mayo went through a similar experimental regimen to reduce the incidence of blood clots and has achieved similarly excellent results. The results of CER were used to develop a protocol that was sensitive to the particular needs and conditions of individual patients—and the result was improved health and lower cost.
Intermountain Healthcare

Intermountain Healthcare is a nonprofit health care provider with 21 hospitals, 140 clinics, over 750 doctors, and 28,000 employees serving Utah and parts of Idaho.

Intermountain providers deliver some 34,000 babies annually. Recognizing that a significant number of births were being induced prior to 39 weeks of pregnancy, Intermountain researchers studied the effects of inducing labor prior to the standard 40-week mark. Believing there to be a relationship between early induction of labor and admissions of newborns to the newborn intensive care facility, doctors were asked whether they saw any such relationship.

From the doctors’ perspective, there was no increase in intensive care admissions among early-induction babies. A study of thousands of patients showed, however, that babies born prior to 39 weeks had a significantly higher probability of requiring intensive care services.

As a result, Intermountain created protocols requiring a physician who advised an induction prior to 39 weeks to justify its medical necessity. The introduction of these protocols resulted in significant improvements in care. The number of Caesarean sections decreased to 1 in 5 (the U.S. average is 1 in 3), the length of time women spent in labor declined, and the number of admissions to newborn intensive care was reduced significantly.

Accompanying these improvements in care were reductions in cost. The shortened labor periods resulted in 45,000 fewer minutes per year women spent in labor. This reduction in the use of hospital resources and personnel translated into some 1,500 additional births at virtually no extra cost. The increase in the number of normal deliveries ($2,600 each) saved significant amounts in comparison to Caesarean sections (average $7,000 each). Overall, Intermountain saves over $10 million per year thanks to these protocols.

In the area of cardiac disease, Intermountain implemented a simple checklist that reminded physicians to prescribe (in most cases) a particular type of heart medication (beta blockers) when patients were discharged. This simple protocol resulted in a reduction in deaths from congestive heart failure of 23% (450 fewer deaths per year), 900 fewer heart disease hospitalizations per year, and savings of $3.5 million/year treating this condition alone.25
According to the Institute of Medicine, the need for comparative effectiveness research is increasing. “The rate with which new interventions are introduced into the medical marketplace is currently outpacing the rate at which information is generated on their effectiveness and the circumstances of best use….Medical care decision-making is now strained…by the growing number of diagnostic and therapeutic options for which evidence is insufficient to make a clear choice.”

Professor Douglas Wood of the Mayo Clinic suggests one example of research that would improve treatment outcomes:

But, let’s think about knee surgery. The question might be: What works the best to relieve pain and improve the functional capacity of patients who are suffering from severe degenerative arthritis of the knee? Is it medication and physical therapy (or other non-surgical approaches), is it some sort of surgery that does not involve a completely new artificial joint, or is it surgery to put in a completely new artificial joint?...

[C]omparative effectiveness... might tell us that it is best to start with medicine and that surgery is most helpful when medicines fail or when there is already evidence from an x-ray of really severe arthritis. It might also tell us that a full joint replacement is actually better than a limited replacement because it lasts longer and gives better pain relief and mobility. So, even though the initial cost of the joint replacement is higher, it is worth it to the patient because of better functional improvement.

According to Dr. James of Intermountain Healthcare, evidence for best therapies only exists for 10-20% of conditions treated, there being no evidence 80-90% of the time. “We desperately need evidence-based protocols for 80% of the patients that walk in the door.”

The need for additional CER was reiterated recently in a report released by the Institute of Medicine of the National Academies entitled Initial National Priorities for Comparative Effectiveness Research. The report recommended the country establish
“a robust CER infrastructure...to sustain CER well into the future,” and listed, in order of priority, the top 100 topics for comparative effectiveness research. Among those in the top 25 were to compare the effectiveness of the different treatments (e.g., assistive listening devices, cochlear implants, electric-acoustic devices, habilitation and rehabilitation methods) for hearing loss in children and adults, compare the effectiveness of strategies for reducing health-care-associated infections, including catheter-associated bloodstream infection, ventilator-associated pneumonia, and surgical site infections, and compare the effectiveness of genetic and biomarker testing in preventing and treating breast, colorectal, prostate, lung, and ovarian cancer.29

A bipartisan proposal for reforming our nation’s health care system by former Senators Howard Baker, Tom Daschle, and Bob Dole recently came to a similar conclusion:

There are considerable research gaps in what we know about the clinical and cost effectiveness of health care treatments and practices, particularly in the area of “personalized medicine,” which studies treatments for subsets of patients based on clinical history, genomics, and other factors. Similarly, there are significant gaps in knowledge about the most effective approaches for payment strategies, benefit features like formulary designs and copayment structures, and information dissemination.30
Federal Support for Comparative Effectiveness Research

Most of the formal research that is done to examine the effects of drugs or medical devices is conducted by the manufacturers of those products in the course of their development. However, there is evidence suggesting that when a company performs the research on its own product, that research tends to indicate that the product is beneficial to health. For example, a comparison of drug studies funded by drug companies to those funded by non-profit groups showed similar results as to the effectiveness of the drugs, but significantly more favorable conclusions—“a 37% gap”—in the drug company funded studies. Similarly, another study found that 37 out of 38 studies with positive results were published while only 14 of 36 were published where the results were unfavorable to the treatment studied; of those 14, 11 suggested that the treatment being studied was more effective than the F.D.A. review actually found.

Not only is privately-funded CER often biased, but it is not in a health corporation’s interest to spend money to do the sort of research that will be used by the entire community as those who have not paid for the research will benefit from it. Thus it is the government and the nonprofit sector that must be counted on to do the research.

The private sector generally will not produce as much research on comparative effectiveness as society would value. The knowledge created by such studies is costly to produce—but once it is produced, it can be disseminated at essentially no additional cost, and charging all users for access to that information is not always feasible. As a result, private insurers and other entities conducting research on comparative effectiveness often stand to capture only a portion of the resulting benefits and therefore do not invest as much in such research as they would if they took into account the benefits to all parties.

A recent article by two physicians in the New England Journal of Medicine supports the government funding of comparative effectiveness research. “The federal initiative will support research that is both more
comprehensive—encompassing many more treatments and conditions, as well as more complete outcome measures—and more relevant to real-world clinical decisions than traditional clinical research.”

“This research is a public good, like highways and clean air. The private sector is no more likely to identify badly mispriced or potentially toxic treatments than it was to spot badly mispriced or potentially toxic products of the banking industry.”
Conclusion

Comparative effectiveness research is research that leads to an understanding of which medical treatments work the best on which patients. The results of comparative effectiveness research make it possible for “doctors and patients to make smart health decisions founded in sound scientific evidence.”

A surprisingly small proportion of “best practices” for treating medical conditions are known. Comparative effectiveness research is necessary to fill in these gaps and to educate physicians as to the most effective treatments, drugs, and medical devices for their patients. Use of these best practices will result in better care for patients and lower costs for the health care system as unnecessary and ineffective treatments are eliminated. Evidence-based protocols that are sensitive to the individual needs of patients are an important tool in improving care.

While part of the problem is attributable to doctors not using the recommended protocols, the fact is that there are no recommended protocols for the great majority of medical conditions. The failure to use evidence-based treatments results in worse health outcomes for patients and increased costs for the health care system. Providers who have instituted evidence-based protocols have experienced significant increases in patient health and reductions in costs. However, such protocols exist only for a small proportion of medical situations.

The objectivity of government and nonprofit institutions position these parties to do the sort of unbiased research that for-profit medical corporations cannot do.

At the end of the McGlynn study quoted earlier the authors concluded: “The deficits we have identified in adherence to recommended processes for basic care pose serious threats to the health of the American public. Strategies to reduce these deficiencies in care are warranted.”
1. Initial National Priorities for Comparative Effectiveness Research, Institute of Medicine of the National Academies, 2009.


3 ARRA, supra, Title VIII, Department of Health and Human Services, Agency for Healthcare Research and Quality.


9 Id. at 2641.


12 http://science.easternblot.net/?p=489.


14 Id. At 2071.

15 Boden, W., O’Rourke, R., “Optimal

16 Id. at 1509-11, 1514.


19 Center for Disease Control, http://www.cdc.gov/nchs/data/hus/hus08.pdf#071.


22 Id. at 3040, 3041.


24 Interview with Douglas L. Wood, M.D., Professor of Medicine, Mayo Clinic, Rochester, MN, June 2009.

25 Information regarding Intermountain from an interview with Brent James, M.D., M. Stat., Executive Director, Intermountain Healthcare Institute for Healthcare Delivery Research, June 2009.


27 Interview with Professor Wood, *supra*.

28 Interview with Dr. James, *supra*.


34 “Research on the Comparative Effectiveness of Medical Treatments,” Congressional Budget Office, *supra*, at 8.


38 McGlynn, E., “The Quality of Health Care Delivered to Adults in the United States,” *supra*, at 2635.