



Society for Medical Decision Making (SMDM)
Background Paper on Comparative Effectiveness Research

May 2009

Mark S. Roberts, MD, MPP
University of Pittsburgh
President

Dave Sugano, DrPH
Novartis Pharmaceuticals Corp.
Vice President

David Paltiel, PhD
Yale University
Secretary-Treasurer

Gillian Sanders, PhD
Duke University
Past President

Kathryn McDonald, MM
Stanford University
President-Elect

Nananda Col, MD, MPP, MPH
Maine Medical Center
Vice President-Elect

Miriam Kuppermann, PhD
University of CA, San Francisco
Secretary-Treasurer-Elect

Board of Trustees:

Ahmed Bayoumi, MD, MSc
St. Michael's Medical Center

Elena Elkin, PhD
Memorial Sloan Kettering
Allison Rosen, MD, MPH, ScD
University of Michigan

Angie Fagerlin, PhD
University of Michigan

Heather Taffet Gold, PhD
Weill Medical College of Cornell
Uwe Siebert, MD, MPH, MSc, ScD
Health and Life Sciences Univ.

Scott Braithwaite, MD
VA CT Healthcare System

Cindy Bryce, PhD
University of Pittsburgh

Bruce Schackman, PhD
Weill Medical College of Cornell

Mark Helfand, MD, MPH
Oregon Health Science University
Editor-in-Chief, Medical Decision Making

Stephen Pauker, MD
Tufts Medical Center
Historian

Jill Metcalf, MBA
Executive Director

SMDM and Comparative Effectiveness Research

- The purpose of comparative effectiveness research is to find out what works best in healthcare in order to help patients and doctors make better healthcare decisions.
- The Society for Medical Decision Making (SMDM) is the leading academic society concerned with making better healthcare decisions.
- Many SMDM members act as liaisons between the scientific community and the health policy community, and have leadership roles in prestigious organizations such as the Institute of Medicine.
- SMDM has expertise on all major methodological aspects of comparative effectiveness research (e.g., evidence synthesis, decision analysis, decision psychology, health economics, cost-effectiveness analysis, mathematical modeling and simulation, and value of information analysis). These complex scientific methods are critically important to asking the right questions and getting the right answers.
- Much like it wouldn't make sense to design a human genome research project without talking to geneticists, it wouldn't make sense to design a comparative effectiveness research portfolio without talking to experts in evidence synthesis, decision sciences, and other relevant methodologies. Jill Metcalf, MBA; *Executive Director*, Society for Medical Decision Making (619-316-7795; jill.metcalf@smdm.org) can assist in identifying SMDM member(s) with the most expertise in a particular area.

What is Comparative Effectiveness Research?

Comparative Effectiveness Research (CER) is research that compares the outcomes of alternative therapies or strategies used to prevent, treat, diagnose, and manage a disease or health condition. The purpose of this type of research is to help patients and physicians make informed health decisions. By providing evidence about which strategies work best for which patients, CER can facilitate better health care choices and thereby improve the quality of health care.¹

Understanding which treatment choices give better or worse results has accounted for only 1.5% of medical research expenditures, less than one out of every thousand dollars spent on health care.² Most of the studies funded today are designed to enable a new treatment to obtain regulatory approval. Such studies usually offer only limited guidance to doctors and patients, often because they compare a new treatment to a sugar pill, placebo, not to the best alternative that the doctor would be considering. Also, these studies evaluate treatments in ideal circumstances (“efficacy”), which often don’t translate well into real-world circumstances (“effectiveness”). In contrast, comparative effectiveness studies compare treatments to their next best alternatives, rather than to placebos, so their results are more useful for informing real-world decisions .

All too often, medical advice is based on limited information and hunches, rather than on rigorous scientific evidence and systematic evaluation. And for some conditions, there may be conflicting information from different sources, with little guidance for physicians and patients who are trying to sift through it to make a good decision.

Recent CER studies have provided valuable information to answer questions like these:

- For a patient who has severe coronary artery disease and must choose between surgery and a less invasive procedure, what is best?³ Is the answer the same if the patient recently had a heart attack?⁴
- Should men be screened for prostate cancer with the PSA blood test?^{5,6} Is the answer the same if they are sick from other conditions?
- How can we choose the best high-blood-pressure drug?⁷ Is the answer the same for patients who are at high risk for heart attacks and strokes?

These are just a few selected examples; the list is as long as the possible situations patients face. Comparative effectiveness research can also be used to compare different hospitals or health systems in which those therapies are delivered (for example, is hospital A or hospital B more effective at treating prostate cancer?)

Right now, it's as if different treatment alternatives are written on a menu in a foreign language. Patients may be able to pronounce their names (for example, *brachytherapy* or *DaVinci surgical technique* for treating prostate cancer), but they can't translate them into terms that are familiar or meaningful to them (for example, changes in survival or quality of life).

CER should help us to get more “bang” for our healthcare buck

While healthcare decision making should be based on many factors, value is one of the important ones. But, unfortunately, we have very little information on the value of different health care services. We only know that, in general, some health care systems vastly underperform others, even when they spend more money. For example, even though health outcomes in the U.S. lag behind most other industrialized countries, the U.S. spends more per person on health than any other country,⁸ with health costs increasing about 2.5 percentage points faster per year than the gross domestic product (GDP), and with federal spending on Medicare and Medicaid possibly reaching 20% of the GDP by 2050 (the same share of the economy that the entire federal budget accounts for today).⁹ Improving the value of healthcare systems will allow us to save more lives and to live more high-quality years, so we get more “bang” for our healthcare buck. CER should provide us with information to strategize about how to get the most value for our healthcare dollar, so we can live longer with better quality-of-life for the money that we spend.

Simply put, value is the amount of additional benefit that a medical treatment gives compared to its additional cost. If a new treatment has a large additional benefit compared to the existing treatment but a small additional cost, then it is high value. If a treatment has a very small additional benefit but a very high additional cost, then it is low value. CER should include measures of benefit that are important to patients, namely quantity and quality of life. Quality of life should be measured using a scale that can be applied broadly, across different diseases and patient groups, so that the value of one treatment can be compared with the value of another treatment. CER is strengthened if it includes measures of costs, considering not just expenses that accrue to any one party (for example, the hospital, the pharmacy, or the insurance company), but also considering all costs that accrue throughout the system, including to the patient and the patient's family, including future saving due to preventive care and better chronic disease management today. While measuring costs is not a mandatory component of CER, it strengthens the usefulness of CER immensely because it provides essential information for estimating value. Otherwise, it would be like

translating the foreign-language menu of healthcare choices into your own language, but then putting “white-out” over the prices.¹⁰

CER should lead to less healthcare rationing, not more

Rationing of health care goes on right now in every healthcare system, and takes many forms (queues, delays in care, denial of coverage, denial of reimbursement, uninsurance, burdensome authorization requirements, high pricing, and other “red tape”). Sometimes it’s out in the open, sometimes it’s hidden, but it’s always there. For example, queues are out-in-the-open rationing mechanisms; burdensome authorization requirements and cost-sharing requirements are hidden rationing mechanisms that may only be apparent to doctors, and sometimes are not apparent to anyone. They go on behind closed doors, and rarely consider value.

CER could shed light on these rationing decisions and allow more people to receive high-value care. It could also reduce the “red tape” and bureaucracy that sometimes gets between patients and their doctors, thereby preventing delays in care. For example, many patients experience unnecessarily long stays in hospitals because their insurers will not pay for low-molecular weight heparin, a medication that is expensive but which also can reduce their hospitalization time and improve their health outcomes. Low-molecular weight heparin is a high-value treatment,¹¹ but one that people too often cannot access because our health system typically focuses on short-term costs rather than long-term value.

CER would translate the health care menu with prices unhidden, and would make the rules transparent so patients could make their own decisions. For example, in the U.S. patients must navigate a foreign-language menu while others watch over their shoulders and attempt to interfere when something is expensive (e.g., saying it’s no longer on the menu, or may take a long time to arrive). CER can make the rules clearer and more fair for everyone.

How CER SHOULD be used

CER should be considered, together with other priorities and goals, to inform decision making by patients, providers, payers, and plans. Some examples of decisions that might be informed are:

- Which therapies works best for which patients?

- How can we improve clinical guidelines?
- How can we provide clinicians with information so they can better provide patient-centered care?
- When is “personalized medicine” most important?
- When should patients have copayments and deductibles waived for a particular health service because its value is so favorable?
- How much is a particular treatment “worth,” based on the amount of benefit that it gives? How can we make sure that patients and health plans do not pay more than this worth?
- Can we pay clinicians and health plans for improving health (e.g., coordinating care), not just for doing procedures and tests?
- Should patients receive a “health dividend” for healthy behaviors that save money for the healthcare system, such as preventing kidney failure in diabetics?
- How do we design future clinical trials to ensure that we can detect improvements that are sufficiently large to offer favorable value?
- What monetary incentives should we give to the highest-quality providers and health systems?

How CER SHOULD NOT be used

- As a sole criterion for denying or awarding care
- As sole criterion for denying or awarding reimbursement to patients or providers
- For hidden rationing of care
- To consider cost in making choices of care or coverage without also considering effectiveness

Summary

CER can help physicians and patients make better decisions about medical care, and to consider value so that we get more “bang” for our healthcare buck. CER is designed to find the best treatment for each patient, and should allow more thoughtful patient-centered care. In contrast, most of the studies funded today are designed to enable a new drug to obtain approval. Such studies usually offer only limited guidance to doctors and patients, often because they compare a new treatment to a sugar pill, not to the best alternative that the doctor would be considering. Real decisions that patients and their families face will often be based on limited evidence and best guesses until researchers focus more on weighing the benefits and harms of available options. CER will help us move the science of health care research in this direction.

The Society for Medical Decision Making (<http://www.smdm.org>) is the leading academic society concerned with making better medical decisions. SMDM members have expertise in all major aspects of CER, especially the complex methods that are essential to asking the right questions and getting the right answers. SMDM supports funding CER because it is a vital investment in achieving health for the nation's population.

TECHNICAL APPENDIX

CER: DEFINITIONS

The Institute of Medicine's definition of CER:¹² The generation and synthesis of evidence that compares the effectiveness of alternative methods to prevent, diagnose, treat, monitor, and improve delivery of care for a clinical condition. The purpose of CER is to assist patients, clinicians, purchasers, and policy makers in making informed health decisions.

Alternative Interventions: Comparators might include systems of care as well as specific interventions to address the prevention, diagnosis, treatment, monitoring, or delivery of care. One comparator could be the current standard of care or usual care.

Study Methods: There are many methods that can be used to conduct CER. A randomized controlled trial between the various treatment options could be conducted, which might provide strong evidence for one therapy's benefit over another, but large, multi-arm trials are expensive, require long periods of time to determine results, and are difficult to update when new technologies appear. Meta-analyses and systematic reviews of existing clinical trials often provide substantial information regarding the relative benefits of different therapies, and may expand understanding of the effects of therapies in certain subgroups, but they require that multiple studies have already been conducted in similar patient populations. Indirect comparisons of separate randomized controlled trials using Bayesian and other methodologies can be useful in the absence of a full meta-analysis. Observational analyses of existing data, (e.g., Framingham study), registries, and administrative or clinical data from electronic medical records, can often provide important insights into the effectiveness of therapies in common use, rather than in the specialized populations that are often enrolled in clinical trials, but they are subject to several biases that require complex statistical methods to mitigate. Finally, there is a growing use of computer simulations and other types of mathematical modeling methods to estimate the differences in treatment effectiveness.

CER: IMPORTANT QUESTIONS

What to study first? We should preferentially study questions for which additional data or analysis is likely to influence the decision, and therefore to confer the highest *value of information* (e.g., the monetarized value of improvements in health outcomes resulting from the better decision-making that is made possible by the study).¹³ We should focus

on healthcare services that are in common practice, generate high costs per-patient-treated or large expenditures in aggregate, employ rapidly changing technologies for which multiple alternative therapies exist for the same problem, or are in clinical, preventive or treatment areas with substantial uncertainty.¹⁴

Which health outcomes to use? In order to estimate value or to permit their results to be used for value estimation in the future, comparative effectiveness studies should include quality-of-life measures that are preference-weighted and gives their results on a unidimensional, interval scale. There are a limited group of quality-of-life instruments that are both easy to use and which satisfy these criteria (EQ-5D, HUI, QWB, SF-6D), however no single measure has ideal properties so that analyzing results according to several different measures is desirable. Comparative effectiveness studies should either employ at least one of these measures, or should employ a measure that has a scientifically validated conversion scale to at least one of these measures.

What cost perspective to use? To consider costs borne by patients as well as those borne by payers, studies should assess costs using a societal perspective, should encompass the longest time horizon that is feasible, and should assess their time-course so that analysts may discount these costs as appropriate (e.g. 3% discount rate). Costs should reflect the opportunity cost of resource allocation rather than charges.¹⁵

How to specify subgroup analyses? Subgroup analyses should be specified *a priori* and should reflect biologically plausible heterogeneity (e.g. age, comorbidity, disease severity), or heterogeneity that may impact groups that have not benefited proportionately from health research in the past (e.g., women, children, certain racial/ethnic groups)

What to use as a threshold for acceptable value? Value thresholds are subject to intense debate,^{16, 17} and vary from time to time and from society to society. That being said, it may be possible to use revealed preferences to make inferences regarding plausible bounds for a particular society's willingness to pay for health benefits, and this can be used to infer a value criterion. For example, the appropriate value threshold in the US is likely to be between \$100,000 and \$300,000 per QALY (quality-adjusted life-year),¹⁸ although the bounds of this range remain controversial.¹⁹

What statistical criterion to use for hypothesis testing? It is customary to mandate greater than a 19-in-20 chance that an improvement is "real" rather than a statistical fluke (i.e., p -value < 0.05). This might be too strict of a standard when treatments are being compared against other treatments, rather than against placebo, particularly if the

treatment is already known to have a favorable side effect and toxicity profile based on prior study. More research is needed to elucidate if and when less strict statistical criteria are appropriate,²⁰ and how this may vary with budget constraints for gathering additional evidence.

What threshold of evidence to require for decision making? It is difficult to specify *a priori* what threshold of evidence is sufficient to influence a decision, since this will be based on how the certitude of evidence compares with the certitude of the decision maker's beliefs in the absence of that evidence. If the decision maker has certitude of belief that greatly exceeds the certitude of the evidence, then the decision maker's post-test probability distribution will not differ much from his pre-test distribution, and the evidence will influence his decision little. If the decision maker has certitude of belief that is greatly exceeded by the certitude of the evidence, then the decision maker's post-test distribution might differ greatly from his pre-test distribution, and the evidence should influence the decision greatly. In order to help decision makers interpret the evidence appropriately, comparative effectiveness studies should express strength of evidence using a simple, transparent, and uniform quality-of-evidence scale (e.g., that used by the United States Preventive Services Task Force).^{21, 22} It is important to note that uncertain decisions with high evidence thresholds will often have high value of information, and are especially suitable candidates for future study.

CER: POLICY IMPLICATIONS

Methods to use CER to inform decision making include the following:

Informing clinical guideline development: Knowing which therapy works best for which patients is the cornerstone of CER, and should make it possible for clinical guidelines to be improved and to incorporate "personalized medicine" (e.g., varying treatment recommendations by comorbidity burden, preferences, and pharmacogenetics).

Figuring out what does not work in health care. Much of our health care budget is spent on services that are not beneficial, and may even be harmful. Investment in CER will allow these services to be identified so they no longer consume valuable resources.

Influence cost-sharing: Low-value health services may be designated for higher cost-sharing schedules, whereas high-value health services may be designated for lower- or no-cost-sharing schedules.⁸ In the situation where a health service is cost-saving, it may be desirable to share this cost-saving with the patient (i.e., a "health dividend").

Influence reimbursement amounts: Reimbursements may be specified so as not to exceed the value of the corresponding health service, given a known magnitude of benefit. For example, if a health service is known to deliver 0.1 additional year of high-quality life as a benefit, the maximum reimbursement for that service would be \$10,000 assuming a value of health benefits of \$100,000 per QALY.

Influence reimbursement criteria: Reimbursement criteria may be specified so as to not exceed the value of the corresponding health service, given a known cost. For example, if a health service is known to cost \$10,000, then reimbursement may be limited to those circumstances in which a patient receives a benefit consistent with 0.1 additional years of high-quality of life, assuming a value of health benefits of \$100,000 per QALY.^{23, 24}

Influence pay for performance criteria: Value estimation can be used to guide incentives for health systems and/or providers for consistently providing services that are known to be high-value. If hospital A provides a high-quality service in 90% of patients and hospital B provides a high-quality service in 95% of patients, knowing the value (i.e., how much is the additional quality “worth”) can inform decisions about the magnitude of pay for performance incentives.

Estimating value of quality differences across health plans or systems of care: Value estimation can be used to estimate the worth of improvements in quality if the resulting increments in quality or quantity of life are known. Suppose Plan A has 80% compliance and Plan B has 60% compliance with an evidence-based colorectal cancer screening program. Because colorectal cancer screening confers 0.25 additional QALYs, on average plan A delivers 0.05 more QALYs than Plan B. Therefore, the value of this quality improvement is worth \geq \$5,000, assuming a value of health benefits of \$100,000 per QALY.

Adherence with principles of value-based insurance design: Value estimation can be used as a more objective measure of whether plans adhere with principles of value-based insurance design.^{8, 25} Based on cost-effectiveness registries and literature reviews, it is possible to construct and maintain an exhaustive list of high-value drugs/services. Then, plans can be compared based on the % of these services that have no applicable copays or deductibles for these services. For example, if Plan A had 95% compliance with value-based insurance design principles (of 20 value-based services, 19 have waived copayments and deductibles) and if Plan B had 50% compliance with value-based insurance design principles (of 20 value-based services, only 10 have waived copayments and deductibles), then Plan A offers more value than

Plan B, and may be more desirable to consumers facing a choice between the two plans.

Estimating return-on-investment (ROI): It is inappropriate to include only financial consequences in an ROI calculation of the costs and benefits of an intervention. From a societal perspective, if an ROI is required, health benefits should be monetarized (e.g., an intervention that delivers 1 high-quality year of additional life delivers at least \$100,000 worth of health benefits) and included in ROI calculations. It is imperative to consider health benefits when estimating return on investment because otherwise ROI calculations just reduce to budgetary-impact calculations, which ignore the benefit of investing in health.²⁶

Informing the design of future clinical trials: If the cost of an intervention is known, a clinical trial can be powered sufficiently to ensure that the incremental effect size is sufficiently large to confer acceptable value given the incremental cost.

References

1. Description of policy options: transforming the health care delivery system: proposals to improve patient care and reduce health care costs. Senate Finance Committee, April 29, 2009.
2. Moses H, Dorsey ER, Matheson DH, Their SO. Financial anatomy of biomedical research. *JAMA* 2005; 294:1333-1342.
3. Serruys PW, Morice MC, Kappetein P, et al. Percutaneous coronary intervention versus coronary artery bypass grafting for severe coronary artery disease. *N Engl J Med* 2009; 360:961-72.
4. Mark DB, Pan W, Clapp-Channing NE, et al. Quality of life after late invasive therapy for occluded arteries. *N Engl J Med* 2009; 360:774-83.
5. Andriole GL, Crawford ED, Grubb RL, et al. Mortality results from a randomized prostate cancer screening trial. *N Engl J Med* 2009; 360:1310-9.
6. Schroder FH, Hugosson J, Roobol MJ, et al. Screening and prostate cancer mortality in a randomized European study. *N Engl J Med* 2009; 360:1320-8.
7. Jamerson K, Wever MA, Bakris GL, et al. Benazepril plus amlodipine or hydrochlorothiazide for hypertension in high-risk patients. *N Engl J Med* 2008; 359:2417-28.
8. Braithwaite RS, Rosen AB. Linking cost-sharing to value: an unrivaled yet unrealized public health opportunity. *Ann Intern Med.* 2007; 146(8): 602-605
9. Orszag PR, Ellis P. The challenge of rising health care costs – a view from the Congressional Budget Office. *N Engl J Med* 2007; 357:1793-1795.
10. Garber AM. A menu without prices. *Ann Intern Med* 2008; 148:964-966.
11. Gould MK, Dembitzer AD, Doyle RL, Hastie TJ, Garber AM. Low-molecular-weight heparins compared with unfractionated heparin for treatment of acute deep venous thrombosis. A meta-analysis of randomized, controlled trials. *Ann Intern Med* 1999; 130:800-9.
12. IOM roundtable on evidence-based medicine: Working group on sustainable capacity. Learning what works best: the nation's need for evidence on comparative effectiveness in health care. September, 2007. Institute of Medicine.
13. Claxton K, Posnett J. An economic approach to clinical trial design and research priority setting. *Health Econ* 1996; 5:513-24.
14. The Society for Medical Decision Making, 2008. "Technology assessment one-pager." The Society for Medical Decision Making urges Congress to include 7 key elements in comparative effectiveness legislation.
15. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. 1996, Oxford University Press; New York, New York

16. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what is it and what that means. *Pharmacoeconomics* 2008; 26:733-44.
17. Culyer A, McCabe C, Briggs A, et al. Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. *J Health Serv Res Policy* 2007; 12:56-58.
18. Braithwaite RS, Meltzer DO, King JT Jr., Leslie D, Roberts MS. What does the value of modern medicine say about the \$50,000 per quality-adjusted life year decision rule? *Med Care* 2008; 46(4):349-356.
19. Hirth RA, Chernew ME, Miller E, Fendrick AM, Weissert WG. Willingness to pay for a quality-adjusted life year: in search of a standard. *Med Decis Making* 2000; 20:332-342.
20. Claxton K. The irrelevance of inference: a decision making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999; 18:341-64.
21. Sawaya GF, Guirguls-Blake J, LeFevre M. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. *Ann Intern Med* 2007; 147:871-5.
22. Braithwaite RS, Roberts MS, Justice AC. Incorporating quality of evidence into decision analytic modeling. *Ann Intern Med* 2007; 146(2):133-141.
23. Garber AM, McClellan MB. Satisfaction guaranteed – “payment by results” for biological agents. *N Engl J Med* 2007; 357:1575-1577.
24. Claxton K, Lindsay AB, Buxton MJ, et al. Value based pricing for NHS drugs: an opportunity not to be missed? *BMJ* 2008; 336:251-254.
25. Fendrick AM, Smith DG, Chernew ME, Shah SN. A benefit-based copay for prescription drugs: patient contribution based on total benefits, not drug acquisition cost. *Am J Manag Care*. 2001;7:861-867.
26. Rosen AB. Grounding coverage in value: a paradigm for linking quality and costs. *Medical Care* 2006; 44:389-391.