



# Comparative Effectiveness Forum

## Executive Summary

November 30, 2006  
Washington, DC

*Sponsored by:*



©2006 Health Industry Forum.  
All rights reserved.



## Table of Contents

Summary	Session Title	Speakers · Panelists	Page
Key Themes	Comparative Effectiveness Forum: Key Themes	<b>Overview</b>	4
Session 1	Comparative Effectiveness: Moving Towards a Consensus for Action	<b>Stuart Altman, PhD</b> <i>Brandeis University</i> <b>Sean Tunis, MD</b> <i>Center for Medical Technology Policy</i>	5
Session 2	Using Evidence to Improve Value: What Do Decision Makers Need?	<b>Jed Weissberg, MD</b> <i>The Permanente Federation</i> Respondent: <b>Peter Bach, MD, MAPP</b> <i>Memorial Sloan-Kettering Cancer Center</i>	7
Session 3	Building Blocks of Comparative Effectiveness	<b>Steven Pearson, MD</b> <i>America's Health Insurance Plans</i> Respondent: <b>Jean Slutsky, MSPH</b> <i>Agency for Healthcare Research and Quality</i>	9
Session 4	Straws, Sticks, or Bricks: How to Build a System of Comparative Effectiveness	<b>Gail Wilensky, PhD</b> <i>Project HOPE</i> Respondents: <b>Jack Rowe, MD</b> <i>Columbia University</i> <b>John Calfee, PhD</b> <i>American Enterprise Institute</i>	12

## Comparative Effectiveness Forum: Key Themes

### Overview

Participants at the invitational Comparative Effectiveness Forum were drawn from many of the key stakeholder groups in health care, including patient advocacy groups, physicians, academia, pharmaceutical and medical device manufacturers, private health plans, and federal researchers and insurers. This diverse group expressed general agreement that an expanded capacity for comparative effectiveness research and analysis should be established in the United States. In addition, such research could be further enhanced and coordinated by a new federal entity, referred to here as a Comparative Effectiveness Board (CEB). The Forum addressed the following key questions: 1) What specific elements of comparative effectiveness should be performed or coordinated by a CEB? 2) Based on the elements included in its role, what are the best options for the organizational structure and funding of a CEB? 3) How can the current agreement and momentum among stakeholders be harnessed to support action that culminates in successful federal legislation?

During the Forum, models were presented to illustrate how more evidence on comparative effectiveness could help support patients, physicians, and payers in their efforts to improve the quality and value of health care. Participants concluded that a CEB could oversee reviews of existing evidence and prioritize topics and funding for new research on comparative effectiveness. In addition, there was general agreement that the reviews and research overseen by the CEB should include measures of both comparative clinical effectiveness and comparative value.

### Context

The Comparative Effectiveness Forum, held in Washington DC on November 30, 2006, brought together leading thinkers from academia, industry, and government to examine in very practical terms how to make broadly available comparative effectiveness research a reality in the U.S. This report summarizes the Forum's four discussion sessions, highlighting key points from each.

### Key Themes

#### ▶ **Enhancing comparative effectiveness in the U.S. requires asking not "should we?" but "how?"**

Participants at the Forum generally agreed to the need for greater capacity and coordination of comparative effectiveness research and analysis in the U.S. Better evidence will be useful to patients in deciding among different treatment options, physicians in treating patients, and payers in making coverage and reimbursement decisions. Greater use of data will lead to better outcomes, more efficient resource use and better value.

One way to enhance comparative effectiveness research and analysis would be the creation of an entity (the CEB) charged with assessing the scope and strength of information on the relative clinical (and possibly cost) effectiveness of alternative health care interventions.

Based on the exchange of ideas presented at the Forum and a review of comparative effectiveness programs in other industrialized countries, basic building blocks were identified as essential functions of such programs.

#### **Generally Supported Functions of the CEB**

1. Prioritize technologies for evaluation.
2. Systematically review existing evidence on comparative clinical effectiveness.
3. Fund studies of comparative effectiveness.
4. Conduct studies of comparative effectiveness.
5. Include cost effectiveness and other value measures as part of analyses of comparative effectiveness.
6. Disseminate findings and coordinate efforts to integrate findings into practice.
7. Establish methodological standards for comparative effectiveness research and analysis.

#### **Not Supported**

8. Create clinical guidelines based on evidence.
9. Make recommendations for coverage/funding.
10. Make coverage decisions.
11. Negotiate prices.

Guidelines were seen as desirable but better aligned with other entities such as medical specialty organizations and academic institutions. Decisions on coverage and prices were not seen as within the scope of a CEB. Instead, clinicians, consumers, and payers would determine how best to use the information available from the CEB.

#### ▶ **Participants agreed that several options for structure, location, and funding of a federal CEB would meet the desired criteria of authority, independence, transparency, and political viability.**

Dr. Gail Wilensky suggested that goals of a CEB include credibility, objectivity, transparency, and expediency. Participants from industry added predictability to these criteria.

Ideas for location of a CEB included linkage to the Institute of Medicine (IOM); creation of a Federally Funded Research and Development Center (FFRDC), and creation of a new entity along the lines of the Federal Reserve Board, Securities and Exchange Commission, or the Federal Telecommunications Commission. Many participants suggested that aspects of comparative effectiveness research could reside within or be commissioned from the Agency for Healthcare Research and Quality (AHRQ), with oversight responsibility resting with the CEB.

There were also several options discussed for funding of the CEB, including general federal revenues or a tax on users of the evidence generated (e.g., health plans, other payers and providers).

All participants agreed with the need to shield this entity from political risk. Both creating the CEB and sustaining it, therefore, requires significant consensus and support among stakeholders.

## Comparative Effectiveness: Moving Towards a Consensus for Action

Speakers: **Stuart Altman, PhD**, *Sol C. Chaikin Professor of National Health Policy, Brandeis University*  
**Sean Tunis, MD, MSc**, *Founder, Center for Medical Technology Policy*

### Overview

Participants generally agreed that comparative effectiveness should be done for many reasons -- to improve decision making by physicians, consumers and payers and deliver the best health value. An entity—referred to throughout this document as a Comparative Effectiveness Board (CEB)—would oversee this activity. The questions to be addressed are practical ones: What would this organization's scope be? Where would it be located? How much funding would be required?

These questions are complex and will require consensus among the many stakeholders within health care. Lessons can be learned from numerous entities which have lacked the funding, support, and structure to survive long term.

### Context

Stuart Altman set the context for the Forum by summarizing the consensus from a previous meeting (on April 11, 2006) and by describing the ultimate questions faced by the U.S. health system.

Sean Tunis provided an overview of comparative effectiveness, summarizing current thinking on the what, why, how, who, where, and when of comparative effectiveness.

### Key Points

▶ **Enhancing comparative effectiveness requires asking not "whether", but "how."**

Professor Altman recounted that the overwhelming conclusion drawn from a previous Health Industry Forum meeting was that the U.S. should definitely engage in comparative effectiveness research and assessment. The key is to agree on how to go about it, where to locate it, and how to move the ball forward.

*"Will we continue to allocate an increasing proportion of our national income to health care?"*

—Stuart Altman, PhD

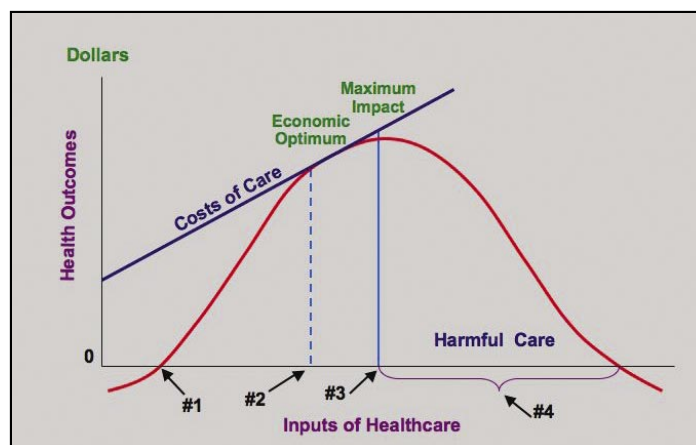
▶ **The value of comparative effectiveness lies in improving how society allocates its resources.**

Speaking as an economist, Stuart Altman framed the ultimate question facing the health system as, "Will we continue to allocate an increasing proportion of our national income to health care?" He suggested that the answer is probably "yes," but that the growth in spending needs to be slowed. Altman pointed out that the growth rate of health spending over the past five to eight years is the steepest in history.

In particular, Altman argued that society must figure out a way to eliminate harmful care and reduce care where the benefits are less than the costs.

Professor Altman presented a chart (below) that compares four combinations of health care costs and outcomes:

- ▶ **Point 1:** Here, the costs of care are low, but so are the outcomes. Everyone would argue for spending more to achieve increased outcomes.
- ▶ **Point 2:** This is the economic optimum where the marginal benefits of care equal the marginal costs.
- ▶ **Point 3:** This is the point where health outcomes are maximized; it is what most physicians and health professionals see as the gold standard of care. However, the costs to achieve these outcomes exceed their economic value to society.
- ▶ **Point 4:** Here, too much care is provided, which not only costs more than is necessary, but is harmful.



In Professor Altman's view, the future policy debate in the U.S. will be whether public funds should be used to pay for services beyond the economic optimum (point #2). In other words, should society seek to maximize the gross benefits of care or the net benefits?

▶ **General agreement exists about the importance of comparative effectiveness and the need for an entity to lead it.**

Sean Tunis stated that there is general agreement among stakeholders that: a new center for comparative effectiveness research (his term for a CEB) should be established in the U.S. to provide information on the relative clinical (and cost) effectiveness of alternative health care interventions and that this center should be funded at a level of \$4 to \$6 billion per year.

*“The health care community believes the time is now to strengthen our ability to use competitiveness effectiveness information to inform health care decisions.”*

—Sean Tunis, MD, MSc

▶ **There has been a great deal of thinking about why such a center is needed, what it would do, and how it would do it.**

Tunis summarized the current research and thinking around the subject of comparative effectiveness, which compares the benefits, risks, and costs of one treatment option to other options. These options are usually drugs, devices, procedures, or diagnostics. “Effectiveness” refers to real world performance (as distinct from “efficacy,” which describes performance in controlled situations).

■ **Why?** Reasons to establish a CEB include: helping payers make more informed coverage and spending decisions; helping patients and clinicians become more informed, cost-conscious decision makers; reducing costs and variations in care; improving quality and safety; sustaining innovation; and improving value.

■ **What else?** In addition to clinical effectiveness and efficacy, this center could: perform Part D drug class comparisons; compare providers on quality and cost; provide condition-based decision guides; and assess geographic variations and perform other health services research.

■ **How?** The center’s major activity could be to perform prospective, head-to-head clinical studies. In addition, the center could undertake observational studies using data from electronic medical records and administrative systems, systematic reviews, and health services research. This differs from the National Institutes of Health (NIH) research which is not necessarily linked to cost.

■ **Who?** There is a great deal of existing research capacity and multiple models. This includes the NIH, the life sciences industry, the Veterans Administration, Centers for Medicare & Medicaid Services, the models used in Europe and Canada, AHRQ, Blue Cross Blue Shield Association, ECRI, ICER, CMTP, and many more. Most of these entities conduct systematic evidence reviews.

▶ **In contemplating where a CEB should be located, there are many ideas and many lessons from the past.**

In Tunis’s view there is not a consensus on where the CEB should be located, but there are valuable lessons from other former undertakings, such as the National Center for Healthcare Technology and the Office of Technology Assessment. Among the lessons are:

- ▶ Small is beautiful but not sturdy. For the CEB to fulfill its intended purposes, it must be sturdy.
- ▶ Creating evidence is safe; making decisions and recommendations is not. Ultimately, decisions and recommendations result in controversy which can jeopardize funding.

- ▶ Transparency, stakeholder input, public accountability, and appeals are necessary to survive.

▶ **There are several “wild cards” that could significantly affect comparative effectiveness initiatives.**

Tunis described wild cards that could have a significant effect on how comparative effectiveness unfolds. These wild cards are:

- ▶ **Payer adoption:** It is not known if and how widely CMS and private payers will adopt a comparative effectiveness approach.
- ▶ **Integrated data:** Not yet known is what can be learned from massively integrated data and how and when broad-based EMR adoption will change the approach to clinical research.
- ▶ **Personalized medicine:** The impact of personalized medicine is unknown. The vision of personalized medicine raises the question of whether large-scale randomized trials will ultimately become obsolete.

## Participant Comments

Following Professor Altman and Dr. Tunis’s presentations, participants raised questions and offered comments. Among the discussion were comments regarding:

- **Clarification of scope.** All parties seemed to agree on a CEB that would cost \$4 to \$6 billion, yet there was not agreement on exactly what the CEB would do. Among the questions that were raised: Would the CEB fund research or just analyze and review? Would the CEB be responsible for dissemination? What would the boundaries for the CEB’s activities be and how would its activities compare with those of the NIH and AHRQ?
- **Human capital.** Given the discussion of the need to increase funding for this type of research, questions arose whether a sufficient number of qualified researchers exist to conduct such studies. If not, these funds will result in bad research. It was suggested to start in a small and focused manner, showing success, creating value, learning, and growing over time. The lack of adequate human capital may be an issue and one option would be to consider providing funds to create the necessary human capital.
- **Measurement risks.** While there was substantial support among participants for the concept of comparative effectiveness research, several risks were also identified:
  - 1) The risk that effectiveness would be judged based on the average results across the entire population, not on efficacy and effectiveness for particular individuals;
  - 2) The risk of a “winner takes all” approach where Option A would be determined to be slightly better than Option B (on average) and, as a result, would be covered while Option B was denied;
  - 3) The risk of making decisions at one moment in time despite the fact that technologies and physicians’ skills change over time. This is particularly important for medical devices where outcomes may depend on the experience and skill of the physician.

## Using Evidence to Improve Value: What Do Decision Makers Need?

- Speaker: **Jed Weissberg, MD**, Associate Executive Director for Quality and Performance Improvement, The Permanente Federation
- Respondent: **Peter Bach, MD, MAPP**, Memorial Sloan-Kettering Cancer Center

### Overview

There is no debate that using evidence improves the care that is delivered. Over the past 10 to 15 years, health systems such as Kaiser Permanente have changed how they set standards and conduct sourcing, making the use of evidence a key part of these processes. Getting physicians to use evidence is more challenging because it requires changing their behavior. Doing so will require not just general evidence, but evidence relevant to sub-segments of patients and evidence that can be applied in real-world settings.

In this drive for improved use of evidence in health care, the patient must not be forgotten as a key decision maker who desires relevant information to make a more informed decision. Patients and patient advocates must be involved in the process of comparative effectiveness so that the evidence generated is usable by all stakeholders, not only by clinicians and payers.

### Context

Dr. Weissberg described how evidence influences decision making at Kaiser Permanente. Dr. Bach responded by highlighting what he sees as some of the goals for comparative effectiveness research and the obstacles faced. Participants then engaged in a discussion about how evidence is currently being used and how it can be better used going forward.

### Key Points

#### ► Providers are increasingly using evidence in sourcing decisions and in setting care standards.

Dr. Weissberg described how Kaiser Permanente's processes have evolved. He reflected how 10 to 15 years ago each hospital would make purchasing and capital allocation decisions based on which department's turn it was to get a new piece of equipment, or which prominent clinician exercised political pull in order to get funds. Non-capital budget decisions were made at lower levels of the organization with little oversight.

That sourcing process has changed dramatically. Now, system-wide purchasing councils oversee decisions in each clinical area (e.g., cardiology, neurology). These councils comprise not just physicians, but also nurses. These councils establish clinical standards and decide which types of equipment, products, and supplies are desired. The decisions now rely very heavily on data, including use of metadata and information from within the delivery system. Decisions also involve the input and collective experience of all of the clinicians on the council.

---

*"Use of meta-analysis is expected in decisions of what to buy."*

—Jed Weissberg, MD

---

These councils recognize that use of data and collective experience are required for the best purchasing decisions. These decisions take into account efficacy, processes, steps, and costs.

#### ► Getting physicians to use clinical data in decision making remains a challenge.

Dr. Bach agreed with Dr. Weissberg's comments on how health systems are increasingly using evidence in purchasing decisions and in standard setting, but remarked that doctors still make too many clinical decisions without data, leading to bad decisions.

---

*"Doctors decide too much without data."*

—Peter Bach, MD, MAPP

---

Dr. Bach emphasized that valuable, useful, actionable comparative effectiveness information can be developed. He believes that a goal for comparative effectiveness research should be to support clinical decision making by identifying specific groups within the population that will derive incremental benefits from specific treatments.

Forum participants mentioned several challenges and obstacles in getting physicians to use evidence in decision making.

- **Having specific evidence.** Often the evidence provided is for entire populations, but physicians may determine that "it doesn't apply to this patient in this situation." Thus, what is needed is evidence that is sub-stratified for particular segments and patients.
- **Translating clinical evidence into practice.** Often clinical research, such as that conducted by the NIH, is developed in very controlled settings but is not easily translatable into practical application at places like community or rural hospitals. For evidence to be truly valuable, it needs to be translated into practice. An issue is that often what researchers want to study—that which is exciting and groundbreaking—is not always what clinicians need.
- **Changing physician behavior.** This is very difficult. Even with sound research, variations in care show that some physicians are providing too much care and need to provide less, while other physicians are providing too little care and need to provide more. Effecting this change is difficult. It extends beyond using evidence and technology; it is changing the culture and environment in organizations.

► **Reimbursement.** In Dr. Wilensky's view, there needs to be a change in mindset away from "all or nothing" reimbursement to differential reimbursement that is based on "for whom and under what circumstances." In this concept, multiple types of treatment would be available to physicians, but the rates for reimbursement would vary based on each patient's particular situation. Important in setting these rates is understanding how much better one treatment option is from another. Wilensky said, "This makes things more complicated, but more useful."

Several participants noted that some professional societies increasingly see it as their professional obligation to collect data and provide evidence in their particular area. Examples given were of the American College of Radiology and the American College of Cardiology. However, Dr. Tunis cautioned that some societies are issuing guidelines not based on any credible data. He noted that having the engagement of a society in creating practice guidelines is beneficial, but it does not equate to having sound empirical evidence.

► **In using evidence to make more informed decisions, the patient must not be forgotten.**

Several participants pointed out that in all of the talk about using evidence to improve clinical decision making, an essential part of the process must be to provide evidence to consumers to help them understand and be actively involved in choosing among different treatment options. As representatives of patients, consumer advocates should be enlisted as part of the process of thinking about and deciding on comparative effectiveness. While it is obvious that consumers should be a critical voice, too often they are forgotten and overlooked.

## Technical Considerations

Participants identified a number of additional issues that are important for considering what a CEB would do.

- **Drugs versus devices.** Assessing the comparative effectiveness of drugs will be different than for medical technologies. For drugs, there are formularies, generics, and a well-established FDA approval process. The process for assessing other technologies is very different.
- **Evaluating combined interventions.** For comparative effectiveness research to have the greatest value, it must look not just at the effectiveness of individual drugs and devices, but at entire interventions. Comparative effectiveness research of entire interventions will provide data on which approaches and methods work better, and not just which drugs perform best. Also, looking at the relative effectiveness of an intervention is important in comparing not just clinical effectiveness, but also costs.
- **Old versus new.** It was mentioned that comparative effectiveness must be conducted for new treatments, but also in assessing existing treatments and possibly even in recommending that some existing treatments be discontinued.



## Building Blocks of Comparative Effectiveness

■ Speaker: **Steven Pearson, MD**, *Senior Fellow, AHIP*

■ Respondent: **Jean Slutsky, MSPH**, *Director, Center for Outcomes and Evidence, Agency for Healthcare Research and Quality*

### Overview

There are core building blocks which are common to organizations that conduct comparative effectiveness research across the world. Many of these building blocks exist in the U.S., but are not centralized or coordinated.

In addition to agreement on the building blocks, also needed are a framework and language for talking about both comparative clinical effectiveness and comparative value; Dr. Pearson presented frame-works for both, including a metric referred to as IVR™ (Integrated Value Rating) which bases a treatment's value on its comparative clinical effectiveness and comparative value.

Important practical considerations in proceeding with a comparative effectiveness initiative include: function (what is to be done); form (where it is to be done); capacity for research and analysis; and linking researchers with decision makers.

### Context

Dr. Pearson presented his thoughts on the core building blocks for comparative effectiveness research and described which of these are used by the health systems in the UK, Australia, and Canada. This provided context for a discussion about which building blocks exist and may be desirable in the U.S.

After Dr. Pearson's presentation, Jean Slutsky shared her perspective on considerations for expanding comparative effectiveness research capacity in the U.S.

### Key Points

#### ▶ Dr. Pearson outlined nine critical building blocks for broadly available comparative effectiveness research.

Dr. Pearson views the key question as one not just of comparative clinical effectiveness, but of comparative value. Having spent considerable time thinking about the functions of a comparative effectiveness program, and having examined the processes used in several other countries, Dr. Pearson offered nine building blocks of a comparative effectiveness research program for discussion:

1. **Prioritize technologies for evaluation:** This would likely be a centralized process of prioritization based on an established set of criteria, such as total system costs.
2. **Systematically review existing evidence:** The idea of reviewing evidence is one of the core functions of any system.

3. **Fund studies of comparative effectiveness:** A potential role may include funding studies where evidence is lacking and deemed necessary.
4. **Conduct studies of comparative effectiveness:** An essential role may be to act as the organization actually responsible for conducting the research that is needed.
5. **Compare cost effectiveness or other value measures:** This element introduces the idea of comparative value as a key building block.
6. **Create clinical guidelines based on evidence:** An output of the comparative research would be to create recommended clinical guidelines.
7. **Make recommendations for coverage/funding:** A function of a comparative effectiveness program could be to provide recommendations regarding what payers cover.
8. **Make coverage decisions:** In some countries, not only is the role of the comparative effectiveness program to review research and offer recommendations, but also to make the decisions for the country regarding what is covered.
9. **Negotiate prices:** Once it is decided what is to be covered, a potential role could be to act as a negotiator of prices.

#### ▶ In contemplating the U.S.'s comparative effectiveness system, an examination of other countries' practices is useful.

The specific building blocks of comparative effectiveness systems vary by country based on each country's health system and culture. Dr. Pearson shared information about the building blocks and practices in the UK, Australia, and Canada.

#### The UK (NICE<sup>1</sup>)

In the UK, NICE prioritizes technologies for evaluation (with \$60 million allocated for this area) and sponsors systematic evidence reviews (building blocks 1 and 2). However, NICE does not fund or conduct clinical studies of comparative research (3 and 4). NICE does compare cost effectiveness and other value measures, only for drugs and devices; not for procedures, and does create clinical guidelines, make recommendations for coverage, and make coverage decisions (5, 6, 7, and 8.) NICE makes recommendations about coverage decisions and those recommendations are always accepted. In fact, there is a requirement that whatever NICE recommends is to be covered. NICE does not negotiate prices.

<sup>1</sup>The National Institute for Health and Clinical Excellence

**Australia (PBAC<sup>2</sup>, PBPA<sup>3</sup>)**

The building blocks that are currently part of the Australian system are blocks 2 (systematic review), 5 (cost effectiveness analysis), 7 (coverage recommendations), and 9 (negotiate prices). The Australian system does not fund or conduct studies, create clinical guidelines, or make coverage decisions. Coverage decisions are made by a government minister who refers to the recommendations made by PBAC. PBPA is responsible for pricing.

A unique aspect of the Australian system is that a drug may be deemed as “cost effective at a specific price,” which serves to spur price negotiation.

**Canada (CADTH<sup>4</sup>, CDR<sup>5</sup>, CEDAC<sup>6</sup>, COMPUS<sup>7</sup>)**

In Canada, there is no process for prioritizing technologies for evaluation because all technologies have to go through an evaluation. All of the other building blocks are incorporated into the Canadian system, with the exception of coverage decisions and negotiating prices, which are left to the provinces. (There is significant variation in the decisions made across the provinces.) Currently, the Canadian system only looks at new treatments, but doesn’t conduct comparative effectiveness research among all treatments.

► **The U.S. system currently has many of the pieces needed for comparative effectiveness research, but these pieces have not been put together.**

In the U.S., most stakeholders agree that building blocks 1-4 are desirable, those being prioritization of technologies for evaluation (1); systematic review of existing evidence (2); and funding and conducting studies of comparative effectiveness (3 and 4).

What is taking place internationally is an important lesson for the U.S., especially the necessity of comparing cost effectiveness and other value measures (block 5).

*“The international example says that you need to look at value.”*

—Steven Pearson, MD

Few U.S. stakeholders are talking realistically about building blocks 6-9 being part of an initial comparative effectiveness system; these being creating evidence-based clinical guidelines (6), making coverage recommendations or decisions (7 and 8), and negotiating prices (9).

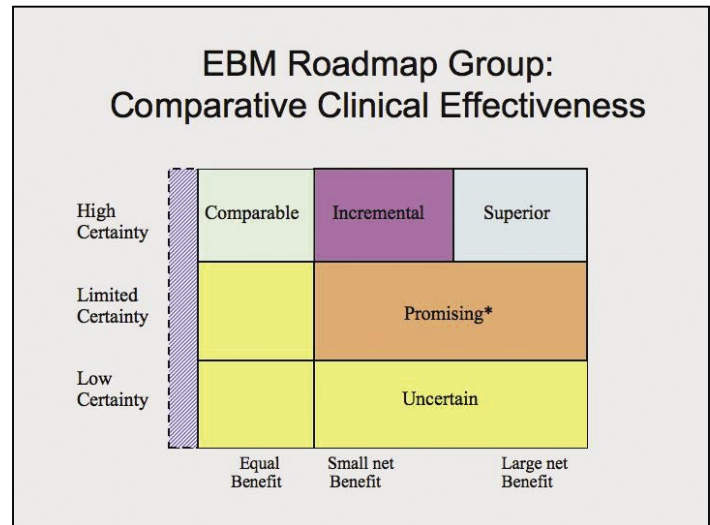
The U.S. already has many of the key elements needed to be part of a comprehensive comparative effectiveness research system. These include organizations that conduct effectiveness

research—including AHRQ, DERP<sup>8</sup>, providers, payers, and private companies that engage in technology assessment. However, there is currently no central coordination of comparative effectiveness research in the U.S.; there are gaps, there is a great deal of waste and inefficiency, and the actual use of evidence is hobbled by lack of a singular credible entity overseeing the process.

► **Recent initiatives provide a framework for thinking about comparative effectiveness.**

Four important initiatives that are addressing aspects of comparative effectiveness are: 1) the IOM EBM, evidence-based medicine Roundtable; 2) the EBM Roadmap Group; 3) the Center for Medical Technology Policy; 4) and the Institute for Clinical and Economic Review (ICER<sup>TM</sup>). These initiatives both deal with clinical comparative effectiveness and assess relative value.

**Comparative clinical effectiveness:** The EBM Roadmap Group has come up with a framework aimed at providing consistency and a common language when talking about comparative clinical effectiveness.



This framework involves thinking jointly about the degree of certainty regarding the evidence surrounding a treatment and the comparative benefit of the treatment. The degree of certainty of the evidence is categorized by three levels:

- 1. Uncertain:** Any treatment where the certainty of the evidence is low is deemed uncertain.
- 2. Limited certainty:** Treatments where the evidence is of limited certainty and where there is some benefit—either small or large compared to current treatments—are deemed as “promising.” Most of what will be evaluated is likely to be seen as “promising.” When the evidence is of limited certainty, it means that there are issues about the quality or generalizability of the data. It might mean that different results have been found for different populations.

<sup>2</sup>Pharmaceutical Benefits Advisory Committee  
<sup>3</sup>Pharmaceutical Benefits Pricing Authority  
<sup>4</sup>Canadian Agency for Drugs and Technologies in Health  
<sup>5</sup>Common Drug Review  
<sup>6</sup>Canadian Expert Drug Advisory Committee  
<sup>7</sup>Canadian Optimal Medication Prescribing and Utilization Service  
<sup>8</sup>Drug Effectiveness Review Project

**3. High certainty:** This means that the quality and strength of the evidence are very high. When the evidence is strong and the benefit of a treatment is equal to that of other treatments, the treatment being assessed is deemed as “comparable.” When the treatment being assessed has a small benefit this treatment is “incremental.” When it has a large benefit it is “superior.” (This means the certainty of the evidence is high and the benefit is large.)

**Integrated Value Rating (IVR)<sup>TM</sup>:** The framework above deals with clinical effectiveness only, not with value. ICER attempts to deal with the value question by creating a separate framework focused on comparative value. ICER is an objective, rigorous, collaborative, transparent model for a public-private organization. ICER aims to test new methods for making technology appraisals accessible and actionable, and provides a metric—the Integrated Value Rating or IVR<sup>TM</sup>—for measuring comparative value.

		Integrated Value Rating (IVR) <sup>TM</sup>		
		COMPARATIVE CLINICAL EFFECTIVENESS		
Superior	A	Ac	Ab	Aa
Incremental	B	Bc	Bb	Ba
Comparable	C	Cc	Cb	Ca
Promising	D	Dc	Db	Da
Uncertain	U	U <sub>-</sub>	U <sub>-</sub>	U <sub>-</sub>
COMPARATIVE VALUE		c Poor	b Reasonable/ Comparable	a Superior

The IVR starts with a treatment’s comparative clinical effectiveness—using the the EBM Roadmap Group framework that deems a treatment as superior, incremental, comparable, promising, or uncertain. Value is then assessed as superior, reasonable/comparable, or poor. (Value is measured from a societal perspective based on the costs and benefits to society.) These measures together yield the IVR, which conveys the comparative clinical effectiveness and the comparative value.

The intent is that a treatment’s IVR would be public information and used to guide action such as reimbursement, pay-for-performance, and formulary decisions.

One participant was uncomfortable with the measure of “value to society” and more interested in the idea of stating “if prices were at a certain level...” Dr. Pearson responded that this is why the terminology is “comparative value” and not “cost effectiveness.” He explained that assessing the comparative value will involve having cost ranges and will entail performing a form of a sensitivity analysis when assessing value.

► **Moving forward with a major comparative effectiveness initiative requires paying attention to practical considerations.**

Jean Slutsky shared her thoughts on a Center for Comparative Effectiveness or CEB. Her views are shaped by AHRQ’s experiences implementing the Effective Health Care Program which produces comparative effectiveness reviews and research.

- **Function should drive form:** In Ms. Slutsky’s view, what the Center is endeavoring to accomplish should drive where the Center is housed and what it looks like.
- **Cultivate capacity:** Adequate capacity is needed both to perform comparative effectiveness research and to interpret and translate findings. These often require different skills and training.
- **Create a bridge to connect researchers and decision makers:** Often researchers do not necessarily design research questions around what policymakers truly need. This creates a disconnect between what research is conducted and what information is needed to make informed decisions. What is needed is a bridge between researchers and health care decision makers so that researchers better address the needs of users.
- **Don’t underestimate the challenge of getting enough patients to participate in clinical studies:** To have adequate patient populations and the ability to sub-stratify, large numbers of patients are needed.

## Straws, Sticks, or Bricks: How to Build a System of Comparative Effectiveness

■ Speaker: **Gail Wilensky, PhD**, *Economist, Senior Fellow, Project HOPE*

■ Respondents: **Jack Rowe, MD**, *Professor, Columbia University*; **John Calfee, PhD**, *Resident Scholar, American Enterprise Institute*

### Overview

Participants see the criteria for a Comparative Effectiveness Board (CEB) as including credibility, objectivity, transparency, and predictability, while creating no delays in comparison to current processes. There was not complete agreement regarding whether comparative effectiveness research should only measure clinical effectiveness or whether it should also involve economic value (with a majority of participants favoring inclusion of economic value).

Most participants envision a center which separates the conducting and oversight of research from the analysis and evaluation of this research. Several individuals envision AHRQ (with additional funding) as managing the research process, with the CEB responsible for evaluation.

Where this CEB would reside is not clear. Ideas that were floated included a new FFRDC<sup>9</sup>; an affiliation with the IOM; or as a new entity with a Federal Reserve-like structure. Whatever the final structure chosen, the CEB will need to withstand changes in political winds, in turn requiring strong support among stakeholders.

Participants were in general agreement that the CEB's focus resides with research and guidance, and that in the U.S., coverage and reimbursement decisions would not be within its domain.

### Context

Dr. Wilensky shared thoughts on the practical criteria for a comparative effectiveness system as well as where such a system might be located, how it would be financed, how its data would be used, and how an implementation process might proceed. Drs. Rowe and Calfee responded by sharing their thoughts on these subjects, and many participants contributed their views.

### Key Points

▶ **Dr. Wilensky argued that the key function of a CEB would be to provide information that enabled payers to make reimbursement decisions.**

Dr. Wilensky distinguishes between “coverage” and “reimbursement.” Decisions regarding whether a treatment should be covered are clinical in nature and based on research about effectiveness and efficacy. In contrast, Dr. Wilensky sees reimbursement as the determination of how much payment for a treatment should be. She believes that reimbursement should be based on a treatment's comparative effectiveness, which means

moving away from a “0 or 1” coverage mentality to a mindset and a process focused on assessing relative/incremental value.

*“The Center needs to get away from a binary, 0, or 1 decision process to a process that decides if there is incremental value.”*

—Gail Wilensky, PhD

Dr. Wilensky believes a co-pay model has merit. It could perhaps be structured where the co-pay is \$0 when a service is supported by evidence as being the most valuable, and the co-pay would increase (perhaps to 100%) when the use of a service is not supported by evidence.

In addition to providing data for reimbursement decisions, the other potential functions that were discussed included:

▶ **Research:** Drs. Wilensky, Rowe, and Calfee see a separation between the roles of conducting/overseeing clinical research and reviewing/evaluating existing evidence for comparative effectiveness. They see AHRQ remaining as the location for oversight of the clinical research, while the CEB would be charged with reviewing the evidence and providing guidance or recommendations.

Dr. Wilensky was emphatic in stating that while comparative effectiveness research in other countries has focused on drugs and devices, it is essential that the U.S. not ignore procedures.

▶ **Comparative physician effectiveness:** Dr. Rowe suggested that there may be more variance in the comparative effectiveness of physicians than in drugs or devices. He believes that invariably an organization named “The Center for Comparative Effectiveness” would be asked to compare physicians' results. He thinks it is important from the outset to state that this is not part of the CEB's mandate.

▶ **Developing human capital:** Dr. Rowe sees a key function of the CEB as “priming the pump” to develop the human capital that will be needed for comparative effectiveness research. He doesn't believe that simply providing funding will result in an adequate supply of people and resources to do the work that is needed. Dr. Wilensky has identified capacity building as a key area of her future research.

▶ **Dissemination:** Many participants view disseminating results as a key function of the CEB. Dr. Rowe argued that the CEB might fund the dissemination process but that this should be done externally by individuals with a different skill set.

<sup>9</sup>Federally Funded Research and Development Center

▶ **Participants see different entities using the information that is generated in different ways.**

Dr. Wilensky sees the role of the CEB as reviewing clinical evidence, with different payers using this information differently. She sees CMS's use of this evidence as extremely important because what CMS does influences other payers.

Dr. Rowe shared a different perspective. He believes that the CEB should go beyond just providing evidence on clinical value and then letting payers make decisions, to providing evidence of economic value. Further, he suggested that recommendations made by the CEB should be followed by Medicare (and maybe Medicaid). Other payers could choose whether to follow or not follow Medicare. Professor Altman noted a parallel to Medicare's use of DRGs for inpatient care; private payers were not mandated to follow this approach, but over time some chose to do so. (Medicare's physician fee schedule has been adapted even more widely.) Those with knowledge of CMS indicated that the agency sees tremendous value in the use of evidence and is already using evidence in a variety of ways. An example is the use of registries.

▶ **The key criteria in establishing a center are credibility, objectivity, transparency, and expediency.**

Dr. Wilensky outlined these criteria based on having visited with multiple stakeholders. Dr. Wilensky emphasized that a new comparative effectiveness board should not introduce delays in bringing new treatments to market. (One participant concurred that the issue of speed is critical, and that the review process could take no longer than six to twelve months.) An idea offered by Dr. Wilensky was that during the comparative effectiveness research process companies could go "at risk."

Participants agreed with these criteria. A participant from the device industry added that an additional essential criterion is "predictability." He also believed that industry would be open to Dr. Wilensky's idea of sharing risk during the research period.

---

*"We will need the system to be predictable; we need to know the rules. We are OK with regulations, but we need to be clear on the path."*

—Device industry representative

---

▶ **Participants differed on where the CEB should be located.**

Dr. Wilensky suggested that to achieve credibility and objectivity the CEB should be outside of government, but close enough to permit financial oversight. For those reasons, Dr. Wilensky offered an intermediate solution of a quasi-governmental agency such as a FFRDC. There are currently 36 FFRDCs including Los Alamos National Laboratory and Lawrence Livermore National Laboratory. Some thought that mismanagement at some FFRDCs might lessen Congressional enthusiasm to create a new one for this purpose.

Dr. Rowe argued that putting the CEB within government would immediately raise the stakes for its long term survival. He suggested that the research funding should go to AHRQ, with other functions based at the IOM, with oversight by other stakeholders. He sees the IOM as a credible, impressive organization that is already focused on evidence-based medicine.

Dr. Calfee agreed with Dr. Rowe on separating out the research but did not like housing the CEB at the IOM. Instead, he favors having multiple entities to create some form of competition and encourage efficiency. He envisioned an entity structured similar to the Federal Reserve with at least half of the members being clinicians.

Uwe Reinhardt, PhD, economist at Princeton University, also supports multiple entities overseeing comparative effectiveness, such as organizations like Robert Wood Johnson.

Dr. Pearson remarked that Dr. Wilensky's focus has been on clinical comparative effectiveness, but the more that cost and comparative value are taken into consideration, the farther outside of the government the CEB should be.

One area of discussion centered on whether the CEB should be affiliated with an existing entity (like the NIH or IOM) to confer legitimacy and credibility, or whether it should be a completely new entity. Both NIH and IOM are credible in the eyes of clinicians, and creating credibility is hard and slow. However, several participants suggested that a completely new entity that was fully supported by stakeholders—and which brought no accompanying baggage—could perhaps quickly achieve even greater credibility. Having multiple entities would not provide the credibility and critical mass needed to generate traction. A patient advocate felt that a newly created entity along the lines of a Federal Reserve model could have a great deal of support among stakeholders.

► **Views on how the CEB should be funded differed, ranging from general revenues to a tax on users (payers).**

Because in Dr. Wilensky's view the CEB is a public good, its funding should come from general revenues, which might entail a tap on the Medicare Trust Fund or other sources.

Dr. Calfee sees the source of funding differently. He agrees that this is a public good but supports Professor Reinhardt's idea of a specific tax which is as close to a market test as possible. By this he means that the users of the research—who are the ones determining what is researched, guiding the research, and benefiting from the research—should be taxed. He likens this to a localized tax assessment. He stressed it is important to keep in mind that the research being conducted is not “for the social good” but is specifically to help stakeholders make economic decisions.

► **Participants agreed that implementation should take place gradually and the structure should protect against politics.**

There was general agreement that it would take time for the CEB to ramp up to full capacity. Participants felt that the CEB should not start out too big, but should start with a more narrow and specific focus; achieve success; and build momentum over time.

Participants were in agreement that the CEB would always be at political risk. Combating this risk can be done through structure (such as an independent Federal Reserve-type structure) but even more importantly, through strong consensus and support among all stakeholders.

Dr. Calfee noted that a system that taxed the users of the research would be less politically risky because those entities paying the tax would see value and would not want the CEB eliminated.

► **Participants identified new building blocks and reached general consensus around which blocks should initially be part of a U.S. comparative effectiveness system.**

At the conclusion of the Forum participants were asked to assess the suggested building blocks that Dr. Pearson had put forth to determine which building blocks should be part of a U.S. system. The results were:

► **New building blocks.** Based on the discussions during the course of the meeting, participants agreed to add two new building blocks. They were: “Disseminate” and “Set standards” (inserted as building blocks 7 and 8). As a result, the complete list of building blocks is now:

1. Prioritize technologies for evaluation.
2. Systematically review existing evidence.
3. Fund studies of comparative effectiveness.
4. Conduct studies of comparative effectiveness.
5. Compare cost effectiveness or other value measures.

6. Create clinical guidelines based on evidence.
7. Disseminate. (new)
8. Set standards. (new)
9. Make recommendations for coverage/funding.
10. Make coverage decisions.
11. Negotiate prices.

► **Participant consensus.** With the 11 building blocks agreed on for discussion, participants then voted on which building blocks should be part of the initial U.S. comparative effectiveness system.

- **Prioritize technologies for evaluation.**  
General support.
  - **Systematically review existing evidence.**  
General support.
  - **Fund studies of comparative effectiveness.**  
There was support for the concept of funding studies where gaps exist, with two different models discussed for actually doing the funding. Splitting off the funding process was favored by most.
  - **Conduct studies of comparative effectiveness.**  
Most supported. There was discussion of the distinction between intramural and extramural studies.
  - **Compare cost effectiveness or other value measures.**  
There was support for performing this task. Most believed this would be done within the same organization; a minority wanted it done by a separate organization.
  - **Create clinical guidelines based on evidence.**  
There was agreement that this task should be done; most supported a separate entity do this work.
  - **Disseminate.**  
General support.
  - **Set standards.**  
General support.
- **Make recommendations for coverage/funding.**  
No broad-based support.
  - **Make coverage decisions.**  
No participants support.
  - **Negotiate prices.**  
No participants support.



