Stakeholder Involvement in Improving Comparative Effectiveness Reviews: AHRQ and the Effective Health Care Program
Research White Paper

Stakeholder Involvement in Improving Comparative Effectiveness Reviews: AHRQ and the Effective Health Care Program

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
http://www.ahrq.gov

Contract No. HHSA 290-2007-10057-I

Prepared by:
AHRQ Effective Health Care Program Product Development Work Group

Authors:
Howard Balshem, M.S.
Pam Curtis, M.S.
Lore Joplin, M.P.A.
Richard A. Justman, M.D.
Alan B. Rosenberg, M.D.

AHRQ Publication No. 11-EHC079-EF
September 2011
Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers; as well as the health care system as a whole by providing important information to help improve health care quality. The reports undergo peer review prior to their release as a final report.

We welcome comments on this Methods Research Project. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by e-mail to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D. Jean Slutsky, P.A., M.S.P.H.
Director Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director, Task Order Officer
Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality
Acknowledgements

Author affiliations are as follows:

Howard Balshem, M.S.
Oregon Evidence-based Practice Center
Oregon Health & Science University
Portland, OR

Pam Curtis, M.S.
Center for Evidence-based Policy
Oregon Health & Science University
Portland, OR

Lore Joplin, M.P.A.
Center for Evidence-based Policy
Oregon Health & Science University
Portland, OR

Richard A. Justman, M.D.
UnitedHealthcare
Minneapolis, MN

Alan B. Rosenberg, M.D.
Wellpoint, Inc.
Chicago, IL

Corresponding author:
Howard Balshem
Oregon Health & Science University
3181 SW Sam Jackson Park Road
Mail Code: BICC
Portland, OR 97239-3098
Phone: 503–220–8262 x54487
Fax: 503–721–1461
Email: balshemh@ohsu.edu

We would also like to thank Susan Stearns, M.B.A., M.A., for her important contributions to the research and feedback on this report.

Members of the AHRQ Product Development Work Group:

Wade Aubrey, M.D.
Senior Clinical and Policy Advisor
Center for Medical Technology Policy

Richard Justman, M.D.
National Medical Director
UnitedHealthcare

Newell McElwee, Pharm.D, M.S.P.H.
Executive Director
U.S. Outcomes Research
Merck & Co., Inc.

John Molina, M.D., J.D.
Medical Director
Arizona Health Care Cost Containment System

Alan Rosenberg, M.D.
Vice President
Medical Policy, Technology Assessment, and Credentialing Programs
Wellpoint, Inc.
President, Anthem UM Services, Inc.
President, Anthem Credentialing Services, Inc.
Stakeholder Involvement in Improving Comparative Effectiveness Reviews: AHRQ and the Effective Health Care Program

Structured Abstract

The Agency for Healthcare Research and Quality (AHRQ) Effective Healthcare (EHC) Program has noted the challenge of decisionmaking when evidence of safety and effectiveness is in dispute and, along with others, has identified the importance of values, preferences, and other contextual factors as part of the decision-making process. Stakeholders have played a significant and important role in identifying ways to make comparative effectiveness reviews more useful to decisionmakers by providing input on the needs of the diverse audience of stakeholders when evidence of safety and effectiveness is weak or uncertain. This paper discusses the valuable input that members of the EHC Program Product Development Workgroup provided regarding several key programmatic and content areas, including: report enhancements designed to support decision making when evidence of safety and effectiveness is limited, weak, or conflicting; improvements to the readability and accessibility of program reports; and enhancements to the future research sections of reports.
Contents

Introduction ........................................................................................................................................1
   Stakeholder Work Group .............................................................................................................1
   Product Development Work Group ..........................................................................................1
      Work Group Goals and Purpose .........................................................................................1
      Work Group Process ...........................................................................................................2

Results .........................................................................................................................................3
   Enhancements to Evidence Reports, Designed To Support Decisionmaking
      When Evidence of Effectiveness is Limited, Weak, or Conflicting ....................................3
   Enhancements to the Future Research Sections of Reports ......................................................5
   Improvements to the Readability and Accessibility of Program Reports .................................6

Conclusions ....................................................................................................................................7

References .....................................................................................................................................8

Tables
   Table 1. Summary of Primary Prevention Trials–Benefits .........................................................4
   Table 2. Summary of Primary Prevention Trials–Harms ............................................................4

Appendixes
   Appendix A. Contextual and Other Factors that Might Influence Coverage Decisions
Introduction

As described by the Institute of Medicine (IOM), “The purpose of comparative effectiveness research is to assist consumers, clinicians, purchasers, and policymakers to make informed decisions that will improve health care at both the individual and population levels.” However, often, when evidence is weak or in dispute, comparative effectiveness reviews (CERs), which synthesize the available evidence of the safety and effectiveness of alternative therapies, may be of limited usefulness to decisionmakers. The Agency for Healthcare Research and Quality (AHRQ), a federal agency charged with improving the quality, safety, efficiency, and effectiveness of health care for all Americans, funds comparative effectiveness research through its Effective Health Care (EHC) Program.

The EHC Program has noted the challenge of decisionmaking when evidence of safety and effectiveness is in dispute and, along with others, has identified the importance of values, preferences, and other contextual factors as part of the decisionmaking process. This paper will describe the important role that stakeholders have played in identifying ways to make CERs more useful to decisionmakers by providing input on the needs of the diverse audience of stakeholders when evidence of safety and effectiveness is weak or uncertain.

Stakeholder Work Group

Authorized in the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, the EHC Program was established to conduct and support research with a focus on outcomes, comparative clinical effectiveness, and appropriateness of pharmaceuticals, medical devices, and other health care services. The program focuses on 14 priority health conditions defined by the Secretary of the U.S. Department of Health and Human Services, and the research is informed by the needs of Medicare, Medicaid, and the Children’s Health Insurance Program (CHIP). The Program includes an 18-member Stakeholder Group, comprising representatives from industry, health plans and insurers, patient advocacy groups, public sector government programs, and purchasers of health care. The role of the Stakeholder Group is to provide input to the Agency and the EHC Program to ensure that its CERs and other evidence reviews are useful to a broad range of decisionmakers, including consumers, clinicians, and policymakers. The Stakeholder Group members are appointed for 2-year terms.

The program’s second Stakeholder Group (serving from October 2007 to February 2010) formed three work groups, whose purpose was to focus more closely on detailed aspects of the EHC Program and to offer more specific input to AHRQ. The three work groups included Program Priorities, with a focus on the processes for nomination and selection of topics for review; Product Development, which focused on issues related to the conduct and reporting of EHC Program research; and Product Dissemination, which addressed issues related to the translation and dissemination of EHC Program products. Each of the work groups reviewed processes and products of the EHC Program and offered suggestions for better achieving the program’s goals. This paper describes the work of the Product Development Work Group.

Product Development Work Group

Work Group Goals and Purpose

The Product Development Work Group consisted of five members of the Stakeholder Group and one or more representatives from AHRQ, and it was staffed by members of the EHC
Program Scientific Resource Center (SRC) and the Center for Evidence-based Policy. Staffing support consisted of methodological and process expertise. The objective of the work group was to provide input to the EHC Program on ways to enhance CER reports so they would be more useful to policymakers and decisionmakers. To achieve this objective, the work group focused on increasing transparency in the following programmatic and content areas:

- Scientific methods;
- Guidance for research;
- Controlling for conflicts of interest;
- Role of consumers in development of scientific reports;
- Quality and consistency of application and implementation;
- Advancing standards for comparative effectiveness reviews;
- Identifying opportunities and new horizons consistent with work group purpose; and
- Improvements to comparative effectiveness reviews to more effectively communicate with diverse stakeholder groups.

**Work Group Process**

The Product Development Work Group met 12 times between January 2008 and February 2010, including six conference calls and six in-person meetings. In addition to participating on conference calls and at face-to-face meetings, work group members reviewed and provided feedback on program products, such as comparative effectiveness reviews and components of those reports. In addition, work group members and staff looked outside of AHRQ and the EHC Program for examples of methods for presenting data regarding the quality and strength of evidence. The work group used these examples to generate discussion and specific suggestions for report enhancements. Meeting summaries and discussion points were disseminated to work group members following each meeting.

In addition to providing reports of discussions and recommendations to the full Stakeholder Group, members from the Product Development Work Group presented at several of the bi-annual meetings of the Evidence-based Practice Centers (EPCs), the AHRQ-funded network of research institutions that conduct CERs and other evidence reviews for the EHC Program.
Results

Members provided specific input regarding several key programmatic and content areas, including:

- Report enhancements designed to support decisionmaking when evidence of safety and effectiveness is limited, weak, or conflicting;
- Improvements to the readability and accessibility of program reports; and
- Enhancements to the future research sections of reports.

Enhancements to Evidence Reports, Designed to Support Decisionmaking When Evidence of Effectiveness is Limited, Weak, or Conflicting

Work group members pointed to the need for information on contextual factors that might influence coverage decisions when there is uncertainty about the safety and effectiveness of a particular therapy or treatment. Contextual factors particularly important to the Work Group when the evidence is low or insufficient included:

- Severity of the disease or condition;
- Whether it is a rare disease, where a robust body of evidence is unlikely to develop within a reasonable period of time;
- Whether there are existing alternative treatments;
- Whether the new treatment is less invasive than existing treatments;
- Patient preferences;
- Already existing community standards with regard to the treatment of a given disease or condition;
- The population affected by the condition;
- Whether future research might influence the decision; and
- Cost effectiveness and pharmaco-economic analyses.

Appendix A was developed as an example of a template that EPCs could use when conducting a CER, to focus their attention on contextual factors that might influence coverage decisions.

The Work Group also suggested improvements to the tables and figures typically produced in EHC Program CERs that would help present situations of low, weak, or conflicting evidence of safety and effectiveness in a way that would more effectively support decisionmaking. After reviewing a variety of table and graphical formats for presenting data on benefits and harms, the Work Group suggested outcomes tables, such as those included below from the review of Medications to Reduce Risk of Primary Breast Cancer in Women, as examples of tables that presented complex information in an easily interpretable and actionable format (Tables 1 and 2).

In addition, the Work Group discussed alternative formats that might more effectively represent complex material and what might be concluded regarding the benefits and harms to patients and consumers of alternative courses of treatment. Examples cited by the group included Consumer Reports Best Buy Drug reports and Institute for Clinical and Economic Review (ICER) reports, such as their review of alternative therapies for prostate cancer. The Work Group also asked that the SRC continue to work with the EPCs on future reports to pilot additional formats for presentation of data.
### Table 1. Summary of primary prevention trials–benefits

<table>
<thead>
<tr>
<th>Major Clinical Outcome</th>
<th>Number of Events Reduced With Medications and Strength of Evidence</th>
<th>Placebo-Controlled Trial—Raloxifene vs. Placebo</th>
<th>Placebo-Controlled Trial—Tamoxifen vs. Placebo</th>
<th>Placebo-Controlled Trial—Raloxifene vs. Tamoxifen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invasive breast cancer</td>
<td>No difference</td>
<td>7 (4, 12) † †+++</td>
<td>9 (4, 14) +++</td>
<td>10 (3, 17) ++</td>
</tr>
<tr>
<td>Estrogen receptor positive</td>
<td>No difference</td>
<td>8 (3, 13) +++</td>
<td>8 (4, 12) +++</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Estrogen receptor negative</td>
<td>No difference</td>
<td>No difference ++</td>
<td>No difference ++</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Noninvasive cancer</td>
<td>No difference</td>
<td>No difference ++</td>
<td>No difference ++</td>
<td>Insufficient</td>
</tr>
<tr>
<td>All-cause death</td>
<td>No difference</td>
<td>No difference +++</td>
<td>No difference +++</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Vertebral fracture</td>
<td>No difference</td>
<td>No difference +</td>
<td>7 (5, 9) +++</td>
<td>44 (25, 61) ++</td>
</tr>
<tr>
<td>Nonvertebral fracture</td>
<td>Insufficient</td>
<td>3 (0.2, 5) ++</td>
<td>No difference +</td>
<td>34 (8, 56) ++</td>
</tr>
</tbody>
</table>

* Study of Raloxifene and Tamoxifen (STAR).
† Number of events reduced compared to placebo per 1,000 women years assuming 5 years of use (95% confidence interval).

### Strength of Evidence Symbols

+++ High, consistent results from numerous (>5) or large definitive trials show a positive protective effect.
++ Moderate, some evidence (3–5 studies) suggests a positive protective effect, but results could be altered by future research.
+ Low, few (<2) trials exist and/or existing trials have inconsistent results regarding protective effect.
Insufficient = data inadequate or not reported.

### Table 2. Summary of primary prevention trials–harms

<table>
<thead>
<tr>
<th>Major Clinical Outcome</th>
<th>Number of Events Increased With Medications and Strength of Evidence</th>
<th>Placebo-Controlled Trial—Raloxifene vs. Placebo</th>
<th>Placebo-Controlled Trial—Tamoxifen vs. Placebo</th>
<th>Placebo-Controlled Trial—Tibolone vs. Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thromboembolic events</td>
<td>6 (2, 10)† More with tamoxifen</td>
<td>4 (2, 9)‡ +++</td>
<td>7 (2, 15) +++</td>
<td>No difference +</td>
</tr>
<tr>
<td>Coronary heart disease</td>
<td>No difference</td>
<td>No difference +++</td>
<td>No difference +++</td>
<td>No difference +</td>
</tr>
<tr>
<td>Stroke</td>
<td>No difference</td>
<td>No difference ++</td>
<td>No difference ++</td>
<td>11 (1, 36) ++</td>
</tr>
<tr>
<td>Endometrial cancer</td>
<td>No difference</td>
<td>4 (1, 10) +++</td>
<td>No difference ++</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Cataracts</td>
<td>13 (5, 21) More with tamoxifen</td>
<td>No difference +</td>
<td>No difference ++</td>
<td>Insufficient</td>
</tr>
</tbody>
</table>

* Study of Raloxifene and Tamoxifen (STAR).
† Number of events increased per 1,000 women years assuming 5 years of use (95% confidence interval).
‡ Number of events increased compared to placebo per 1,000 women years assuming 5 years of use (95% confidence interval).

### Strength of Evidence Symbols

+++ High, consistent results from numerous (>5) or large definitive trials show a harmful effect.
++ Moderate, some evidence (3–5 studies) suggests a harmful effect, but results could be altered by future research.
+ Low, few (<2) trials exist and/or existing trials have inconsistent results regarding harmful effect.
Insufficient = data inadequate or not reported.
Enhancements to the Future Research Sections of Reports

In addition to identifying ways to better present information on existing studies, the Work Group suggested that the future research sections of CERs could provide more comprehensive information on both the likelihood that future research might produce evidence that would influence current practice, and more detailed and specific information about the kind of research required.

The work group suggested that the AHRQ Technical Brief on Particle Beam Radiation Therapies for Cancer\(^9\) provided an example of some of the limitations of current reports. Particle beam therapy (PBT) is a focused radiation therapy used to treat certain arteriovenous malformations and a variety of cancers. It is more focused than standard external-beam radiation therapy, so theoretically its use can spare nearby healthy tissue. On the other hand, it is also not known whether by focusing on a narrower field, PBT may, in fact, miss some occult cancer cells that might reside in otherwise apparently healthy tissue. However, there have been few head-to-head studies comparing PBT with other forms of external-beam radiation therapy. The AHRQ report concluded that preferential use of PBT is not yet supported by clinical evidence.\(^9\)

Various payers define alternative criteria for the medical necessity of PBT including, but not limited to, treatment of arteriovenous malformations and melanoma of the uveal tract, treatment of intracranial and skull base tumors, spinal cord tumors, and prostate cancer. In the context of the uncertainty of the evidence, this diversity of benefit decisions is understandable. However, for patients, consumers, clinicians, and policymakers, the AHRQ review of the published literature left significant unanswered questions regarding this technology including:

- What is the evidence that PBT works at all?
- In the absence of head-to-head trials, how will we accumulate evidence about PBT?
- What type of evidence is necessary to establish the relative benefits and harms of PBT?
- What type of evidence is necessary to establish its efficacy in improving clinically significant outcomes?
- How does PBT compare to other types of focused radiation, e.g., intensity-modulated radiation therapy (IMRT)?

Elaborating on the role of future research, the Work Group recommended that such research should follow the principles of comparative effectiveness research,\(^10\) build on the guidance described in the IOM report on comparative effectiveness research,\(^1\) address the social importance of the research, recognize the diverse audience with interest in the research, and provide sufficient granularity to be actionable. Recommendations included the need to address the societal importance of future research and the need for greater specificity regarding the patients, interventions, comparators, outcomes, timing, and settings (PICOTS) that would increase the likelihood that future research would fill the gaps in current knowledge. These recommendations played an important part in the EHC Program's decision to initiate a new series of Future Research Needs reports,\(^11\) and will continue to inform the content of those reports as they are developed and evolve.

The Stakeholder Group also emphasized the importance of identifying and prioritizing research in the context of its impact on patients and on broad societal interests, rather than relying only on traditional prioritization criteria, e.g., disease prevalence, burden of illness, etc., or on areas of specific interest to researchers. This was a significant focus for the group and was considered a central reason why there remains so much uncertainty and inadequate evidence on key clinical decision points for patients on subjects reviewed in CERs.
The Stakeholder Group also expressed the concern that these reports and the research gaps identified should adequately inform and help prioritize future government research funding by AHRQ, the National Institutes of Health, and other government funding sources related to taxpayer funded research activities.

**Improvements to the Readability and Accessibility of Program Reports**

Finally, members of the Work Group also noted that the program reports can be narrowly technical, and, therefore, frequently not relevant to those responsible for setting policy. After careful review and consideration of six sample CERs, members of the Product Development Work Group made detailed suggestions to improve the readability of EHC reports. These included:

- Crosswalking the Executive Summary with related sections in the full report, and including information from the Conclusion and Future Research sections in the Executive Summary;
- Increasing the use of graphics and diagrams;
- Using author names instead of reference numbers in citations;
- Reducing the use of acronyms and including an acronym glossary;
- Adding a summary introduction and conclusion for each section of the report highlighting the main “take home” messages of that section;
- Highlighting key information that physicians can provide to patients; and
- Providing that information in a way that can be translated into documents that can be easily understood by patients and by nonclinical policymakers.
Conclusions

While acknowledging the excellence of EHC Program CERs, the Product Development Work Group concluded that they are frequently insufficient to address the questions important to the stakeholders who are the intended audience for those reports. They suggested the need for reports to go beyond conclusions of insufficient evidence and statements that future research should address these questions. Rather, stakeholders need evidence on:

- What evidence is available, and what can be learned from it?
- What type of evidence will tell us when and whether a service is safe or harmful and effective or ineffective for significant clinical outcomes?
- What can we learn from evidence that falls short of a randomized controlled trial?
- What role can registries play in telling us whether a treatment is safe?
- What evidence will identify subpopulations most likely to benefit from its use?
- What type of evidence will demonstrate short-term effectiveness?
- What type of evidence is necessary to demonstrate long-term effectiveness?
- When is shorter term evidence appropriate (e.g., if no obvious harm is identified), and under what circumstances should longer term evidence be required (e.g., should implanted devices require longer term evidence than procedures without an implanted device).
- If a service meets an unmet need, is safe, and appears to be effective in the short term, under what circumstances, if any, should consumers, physicians and policymakers make decisions based upon this incomplete evidence?

It is important to acknowledge the challenges faced and contributions made to effective collaboration by both the EHC Program and the Stakeholder Work Groups, and in particular to note the significant investment of time and the important contributions of the Stakeholders. Skilled facilitation that provided the necessary understanding of the Program’s needs as well as continuity between Work Group meetings was essential to the success of that collaboration. While Stakeholders sometimes made suggestions that AHRQ could not implement due to program limitations, the Stakeholder Group members who served on the Product Development Work Group provided AHRQ with thoughtful, specific, and useful input on how to improve and enhance research reports and increase their usefulness to decision makers. In addition, EPC directors and their staff valued hearing suggestions from high-level decision-makers about how to improve the usability of the final reports. Work group members helped the Program to better understand that decisionmakers need to make a decision regardless of the availability or quality of the evidence, and provided specific input designed to enhance the reports to better inform those decisions when limited or lower quality evidence is all that is available.

The work of the Product Development Work Group and its suggestions for improving and enhancing CERs provides a foundation and example of stakeholder input to a federal agency that is charged with the conduct of comparative effectiveness research. As additional federal resources are invested in conducting comparative effectiveness research, the input by decisionmakers to AHRQ and other federal agencies becomes increasingly relevant so that the resulting research reviews meet their needs and address the priority conditions and populations as defined by the original legislation.
References


Appendix A. Contextual and Other Factors That Might Influence Coverage Decisions

STRENGTH OF THE BODY OF EVIDENCE

- Risk of bias of study design: randomized versus observational studies or other types of evidence
- Description of comparators
  - Standard of care
  - Timeliness of data
  - Comparability of populations
- Retrospective studies of survival
- Prospective studies of survival
- Direct or indirect comparisons
- Applicability or generalizability of evidence (i.e., description of interventions, populations, context)
- Clinical relevance/importance of measured outcomes

SUBPOPULATIONS

- Evidence of variations in safety and effectiveness in defined subpopulations

SPECIFIC INFORMATION WHEN THERE IS LOW OR INSUFFICIENT STRENGTH OF EVIDENCE

- Explanation of rating
- Potential impact of making wrong decision
- Will collecting more data be of value
  - What would new research look like
  - Urgency of new research
- What can be said about the strength of the evidence of benefits and harms

INFLUENCING FACTORS AND CONTEXTUAL INFORMATION

- Clinical considerations
  - Life-threatening condition
  - Debilitating or severe disease
  - Rare disease
- Relationship to treatment alternatives
  - No existing alternative treatment
  - Ease of use
  - Improved access
  - Cost compared to alternative therapies
  - Less invasive procedure
  - Patient preferences
- Additional available information
  - Existing evidence-based clinical guidelines
  - Anecdotal evidence of current practice
- Expert opinion
- Advocacy positions/perspectives
- Required training
- Facility requirements
- Long-term versus short-term costs
- How to measure cost-effectiveness/when is QALY analysis not applicable

- Other value propositions
  - How would adoption affect use of other health services?
  - How would adoption affect use of other more invasive, harmful or expensive treatments?
  - How would adoption affect use of treatments by different types of providers that treat the same condition?

- Harms

ONGOING AND FUTURE RESEARCH
- Description of current research that may inform decisionmaking
- How should additional research/resources be pursued? (develop research urgent criteria)
  - Have decisionmakers already determined whether or not to cover? If yes, on what basis was the decision made?
  - Develop criteria for prioritization and prioritize future research needs, i.e., top 10

DECISION AIDS
- Include grid that displays continuum of strength of the evidence as it relates to potential coverage decisions
- Conditions of coverage
  - Medical management
  - Differential benefit design based upon apparent clinical value
- Decision tree