



Community Forum Deliberative Methods Demonstration: Evaluating Effectiveness and Eliciting Public Views on Use of Evidence

Executive Summary



Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Community Forum Deliberative Methods Demonstration: Evaluating Effectiveness and Eliciting Public Views on Use of Evidence

Executive Summary

Prepared for:

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

Contract No. 290-2010-00005

Prepared by:

American Institutes for Research
Washington, DC

AHRQ Publication No. 14-EHC007-EF

November 2013

Revised April 2014

The Agency for Healthcare Research and Quality's (AHRQ) Effective Health Care Program conducts and supports research focused on the outcomes, effectiveness, comparative clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services. More information on the Effective Health Care Program and electronic copies of this report can be found at www.effectivehealthcare.ahrq.gov.

The Community Forum was initially funded under the American Recovery and Reinvestment Act (P.L. 111-5) to expand public and stakeholder engagement in AHRQ-supported comparative effectiveness research programs. This final report was developed with ARRA funding, by the American Institutes for Research, under AHRQ Contract No. 290-2010-00005.

The AHRQ Task Order Officer for this project was Joanna Siegel, Sc.D. The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ or the U.S. Department of Health and Human Services. Therefore, no statement in this report should be construed as an official position of AHRQ or the U.S. Department of Health and Human Services.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials that are clearly noted in the document. Further reproduction of those copyrighted materials is prohibited without the specific permission of copyright holders.

People using assistive technology may not be able to fully access information in this report. For assistance contact EffectiveHealthCare@ahrq.hhs.gov.

James Fishkin at the Center for Deliberative Democracy, Stanford University, was the original developer of "Deliberative Polling[®]," one of the methods included in this study. No other interest or affiliation is noted among other members of the study investigator team.

Suggested Citation:

Carman KL, Maurer M, Mallery C, Wang G, Garfinkel S, Richmond J, Gilmore D, Windham A, Yang M, Mangrum R, Ginsburg M, Sofaer S, Fernandez J, Gold M, Pathak-Sen E, Davies T, Siu A, Fishkin J, Rosenberg M, Fratto A. Community Forum Deliberative Methods Demonstration: Evaluating Effectiveness and Eliciting Public Views on Use of Evidence. (Prepared by the American Institutes for Research under Contract No. 290-2010-00005.) AHRQ Publication No. 14-EHC007-EF. Rockville, MD: Agency for Healthcare Research and Quality; November 2013, revised April 2014. www.effectivehealthcare.ahrq.gov.

Acknowledgments

The American Institutes for Research (AIR) would like to acknowledge the U.S. Department of Health and Human Services' Agency for Healthcare Research and Quality for sponsoring this project. We also thank our Project Officer, Dr. Joanna Siegel, for her leadership and guidance.

We extend gratitude to our multidisciplinary team of organizations and individuals for their expertise, input, and commitment to this 3-year initiative to improve and expand public and stakeholder engagement in the Agency's Effective Health Care (EHC) Program. Below, we list individuals who contributed to the design, implementation, and/or evaluation of the Community Forum Deliberative Methods Demonstration.

American Institutes for Research (in alphabetical order)

Kristin Carman, Ph.D., Project Director

Amar Beeharry

Matthew Borda, M.A.

Johannes Bos, Ph.D.

Graciela Castillo, M.P.H.

Michael Cohen, Ph.D.

Sandy Eyster, Ph.D.

Jessica Fernandez

Kirsten Firminger, Ph.D.

Allison Fratto

Samuel Mintzer Fuchs

Steven Garfinkel, Ph.D.

Dierdre Gilmore, M.A.

Susan Heil, Ph.D.

Jessica Heeringa, M.P.H.

Kyosin Kang

Chelsea Kelleher

Laurel Koester, M.P.H.

Elena Liroff

Coretta Mallery, Ph.D.

Rikki Mangrum, M.L.S.

Diane Martinez, M.P.H.

Regin Mathew

Maureen Maurer, M.P.H.

Marilyn Moon, Ph.D.

Peter Morrison

HarmoniJoie Noel, Ph.D.

Alex Ortiz

Ushma Patel, M.S.P.H.

Jennifer Richmond

Simone Robers, M.A.

Mark Rosenberg

Herschel Sanders

Karthik Shyam, M.P.P.

Adair Smith

Andrew (Kip) Thomson

Grace Wang, Ph.D., M.P.H.

Channet Williams

Amy Windham, Ph.D.

Thomas Workman, Ph.D.

Manshu Yang, Ph.D.

Jiani Yu

Center for Healthcare Decisions

Fern Alexander, Ph.D.

Tamar Foster, M.P.P.A.

Marge Ginsburg, M.P.H.

Kathy Glasmire

Center for Deliberative Democracy, Stanford University

James Fishkin, Ph.D.

Alice Siu Chao, Ph.D.

Kathleen Giles, M.A.

Center for the Study of Language and Information, Stanford University

Noam Ben-Avi
Todd Davies, Ph.D.
Michael Huang
Mike Mintz, M.S.
Kenneth Taylor, Ph.D.
Jimmy Tobin

Consultants

Marthe Gold, M.D., M.P.H., City College of New York
Ela Pathak-Sen, M.B.A., Commotion
Shoshanna Sofaer, Dr.P.H., Baruch College, City University of New York

Technical Expert Panel

Julia Abelson, Ph.D., McMaster University
Stirling Bryan, Ph.D., University of British Columbia
Tim Carey, M.D., M.P.H., University of North Carolina–Chapel Hill
Scott Y. Kim, M.D., Ph.D., University of Michigan
Peter Muhlberger, Ph.D., Texas Tech University
David Ropeik, M.A., Harvard University
Mark E. Warren, Ph.D., University of British Columbia

Expert Presenters for the Deliberative Demonstration

Mark Holmes, Ph.D., University of North Carolina–Chapel Hill
David Hyman, M.D., J.D., University of Illinois
Anthony Iton, M.D., J.D., M.P.H., The California Endowment
Lawrence Kleinman, M.D., M.P.H., Harvard
Christine Mitchell, R.N., M.S., MTS, FAAN, Children’s Hospital Boston
John Morton, M.D., M.P.H., M.H.A., FACS, FASMBS, Stanford University
Ira Moscovice, Ph.D., University of Minnesota
Rita Redberg, FACC, M.D., M.Sc., University of California, San Francisco
Kyu Rhee, M.D., M.P.P., FAAP, FACP, IBM
James Sabin, M.D., Harvard University Medical School
Lisa Schwartz, M.D., M.S., Dartmouth
Mark Tonelli, M.D., M.A., University of Washington
Glennah Trochet, M.D., Sacramento County
Steven Woloshin, M.D., M.S., Dartmouth

Executive Summary

Introduction

The Agency for Healthcare Research and Quality (AHRQ) Community Forum, initiated under the American Recovery and Reinvestment Act (ARRA), has as its goal to improve and expand public and stakeholder engagement in AHRQ's Effective Health Care Program. A primary area of focus for the Community Forum is to advance methods for obtaining input from the general public.

This report describes the results of the Deliberative Methods Demonstration, a randomized controlled trial comparing deliberative methods with one another and with a control intervention. The primary aims of the Demonstration were to:

- **Inform AHRQ research programs on public views regarding the use of research evidence in health care decisionmaking** by obtaining informed public input on questions regarding appropriate and acceptable ways to use evidence that are central to the mission of AHRQ's research programs.
- **Expand the evidence base on public deliberation** by evaluating whether public deliberation is an effective and useful way to obtain informed public input for U.S. health care research, as well as identifying a feasible set of choices among deliberative methods.

What Is Public Deliberation and How Has Its Effectiveness Been Evaluated?

Public deliberation methods provide opportunities to obtain informed perspectives on complex topics that are value laden and that lack simple technical solutions. On such issues, public input on what underlying values should be considered, potential tradeoffs in values, and potential solutions and their likely uptake or resistance are important considerations in developing programs or policies.

Deliberative methods are a distinct approach to obtaining public input. In public deliberation members of the public are convened to obtain input about—and meaningful insights into—how people think about a topic when they are informed. Thus, information obtained through public deliberation differs from that collected through surveys or focus groups, which generally obtain more top-of-mind—that is, initial and more intuitive—responses and reactions. In deliberative sessions, participants receive information that is intentionally neutral and respectful of the full range of underlying values, experiences, and possible perspectives. They are encouraged to discuss, learn from others, and examine and refine their own views.

Although considerable theoretical and case-study literature endorses the value of public deliberation, little empirical research has been conducted about its effectiveness.¹ In the research that has been done, effectiveness has been defined as:

- **The quality of deliberative experience or discourse.** Using participant self-reports, researchers' observations, or reviews of session transcripts, these measures typically assess levels of equal participation, active participation, opportunity for adequate discussion, respect for the opinions of others, and awareness of different perspectives.

- **Changes in participants’ knowledge or attitudes about the deliberative topic.** A core goal of deliberative methods is informed input, and a core assumption is that information and discussion may alter the views of participants. Thus, typically using pre and post surveys, these measures assess the effect of the deliberation on the participants’ knowledge, attitudes, perspectives, values, beliefs, opinions, or policy preferences on the deliberative topics.
- **Changes in participants’ empathy and concern for issues affecting the community at large.** Using pre and post surveys, a number of studies have assessed the effect of deliberation on civic engagement and capacity, engagement in the political process, sense of self-efficacy, sense of empowerment, political efficacy and solidarity, and anticipated post-meeting activity related to deliberation issues.
- **Impact on decisions by sponsoring agency.** Ultimately, deliberation obtains information that can influence decisions. Measurement constructs include the effect of public input on specific laws, policies, or practices and on decisionmakers’ intentions to act on the results of deliberation. These constructs are usually assessed through case studies or surveys of decisionmakers who may use the findings from the deliberation.

Few well-designed comparative studies of deliberative methods or their alternatives have been conducted.

Deliberative Methods Demonstration Description

Between August and November 2012, we conducted a five-arm randomized controlled trial to examine the effectiveness of public deliberation and to compare alternative approaches. Participants were assigned to one of four deliberative methods or a control group. The project convened 76 groups in four locations: Chicago, IL; Sacramento, CA; Silver Spring, MD; and Durham, NC. We selected locations that made it easier to recruit a diverse sample in terms of racial, ethnic, and sociodemographic background, with specific attention to ensuring inclusion of members of three AHRQ priority populations: Hispanics, African-American women, and the elderly.

Deliberative Topic

Across all methods, the Deliberative Methods Demonstration elicited public input on the use of research evidence in health care decisionmaking. We posed the following deliberative question to all participants:

Should individual patients and/or their doctors be able to make any health decisions no matter what the evidence of medical effectiveness shows, or should society ever specify some boundaries for these decisions?

This question was appropriate for deliberation for several reasons. First, the use of evidence in decisionmaking relates directly to AHRQ’s support of research that helps people make more informed decisions and improves the quality of health care services. As such, public input on this question had the opportunity to make valuable contributions to the AHRQ program. Second, the question required participants not only to understand how evidence is generated and used, but also to discuss difficult tradeoffs concerning the impact on individuals and communities when evidence is or is not applied in medical decisions. Finally, responses to the question would elicit

the public's values around whether patients and physicians have a social responsibility to make evidence-based health care decisions.

Prior to their participation, all participants received the *Preparing for the Community Forum* booklet, which described the overall purpose of the project and what to expect (Appendix B). It also gave definitions and facts on medical research and medical evidence, quality health care, and comparative effectiveness research. Information on rising health care costs and who pays for health care was included to provide context for the discussions. We did not provide information on rules, guidelines, or any other types of boundaries in health care; rather, we allowed interpretations and discussions of boundaries to arise spontaneously.

To help participants grapple with a complex topic and a fairly abstract question, we developed specific case studies to provide context for each deliberation (Appendix C). These were:

- *Comparing Hospital Quality*
- *Upper Respiratory Infection in Children: Antibiotics Versus Symptom Treatment*
- *Obesity Management: Comparing Treatment and Prevention*
- *Heart Disease Treatment: Comparing Medicines Only and Stents Plus Medicine*
- *Comparing Approaches To Preventing Illness: A Fictional Case*

All methods used the case study on comparing hospital quality, and two methods used additional case studies.

Deliberative Methods

We selected four distinct types of deliberative methods that have been used in prior public deliberations and reflect important differences in implementation: number of participants, session length, mode of interaction, and use of content experts. We refined each type of deliberative method to ensure that all methods included necessary components of successful deliberation identified in our literature review, while retaining the methods' core distinctiveness.

Brief Citizens' Deliberation (BCD): In this method, 12 participants met in person once for 2 hours. A single facilitator and a single note-taker supported these groups. Facilitation was active, designed to encourage attention to the tensions among social values, ethical principles, and the individual versus societal perspective. Participants discussed the hospital quality case study. No expert presentations were included in this method. We held 24 BCD groups, 6 at each location.

Community Deliberation (CD): This method involved two deliberative sessions, each 2.5 hours long, 1 week apart, for each group of about 12 participants. In the first week, participants discussed the upper respiratory infection (URI) case study. During the week between in-person sessions, participants interacted through an online discussion board. In the online setting, two experts provided statements regarding the URI case study, answered participants' questions, and asked questions of their own. At the second in-person session, participants completed discussion of the URI case study and went on to discuss the hospital quality case study. A single facilitator and a single note-taker supported these groups. During the in-person sessions, facilitation was active, as described for BCD above. We held 24 CD groups, 6 at each location.

Online Deliberative Polling® (ODP): In this method, each group convened online four times, once per week over a period of 4 weeks, using one case study (hospital quality). Each meeting was a 1.25-hour online session, during which about 12 participants engaged in discussions via a

dedicated Web site and Internet-based audio conferencing. Student facilitators with no prior experience in facilitation or health care moderated these groups; they were trained to intervene as little as possible during discussions, while still attempting to ensure consideration of the competing arguments in the reading materials. This facilitation style was put in place in order to maintain the neutrality of the moderator. During the first two sessions, participants began exploring issues about hospital quality. Following discussion, the groups had the opportunity to generate questions to be addressed offline by a panel of three experts. The panelists' responses were played back to participants during the third session and served as a basis for further conversation in the final session. We held 24 ODP groups.

Citizens' Panel (CP): CP involved 2.5 days of deliberation. There were 24 to 30 participants in each group. All five case studies were used. Seven experts were linked to the group through Skype[®] at key points during the session to provide additional information and different points of view on the case studies and issues related to the deliberative question. A clinical expert, who was also a member of the research team, presented on comparative effectiveness research and addressed questions from participants. Three facilitators and a note-taker supported these groups. This method permitted the use of smaller breakout groups moderated by a facilitator, as well as an open space in which participants could interact without facilitation. Facilitation in this method was active, as described for BCD and CD above. We held four CP groups, one at each location.

Reading Materials Only Control Group: Participants assigned to the control intervention received educational materials via an email link. Materials included the same background booklet provided to the deliberative groups, *Preparing for the Community Forum*, as well as three of the case studies: hospital quality, URI, and obesity management. We chose three of the five case studies to present to the control (a midway point between the other methods, which received between one and five case studies). Participants did not convene in groups to deliberate. We estimated an hour of reading time.

Study Sample

Of the 1,774 participants recruited from the four locations, 961 participants took part in a deliberative method and 377 participants were a part of the reading materials only control (an overall show rate of 75 percent). The study sample was diverse and reflected each location's population in terms of sex, age, race, and ethnicity based on U.S. Census Bureau estimates, but had a larger percentage of people with at least some college education.

Findings

Public Views About Use of Evidence in Health Care Decisionmaking

To address our first aim, we conducted a thematic qualitative analysis of transcripts from the 76 deliberative groups to summarize how participants responded to the overall deliberative question. The research questions for the thematic analysis focused on three main topics related to the overarching deliberative question: (1) circumstances participants specify for restricting decisionmaking, (2) situations affecting how participants perceive those circumstances, and (3) the social values exhibited during deliberation.

When asked the overarching question, participants first focused on the concept of boundaries. Many of them initially interpreted boundaries as compulsory rules that limited choices and

allowed no exceptions, and most reacted negatively. Participants also questioned what was meant by “society.” They initially defined society as the government or a health insurance company—perceiving both types of organizations as enforcers of boundaries in health care. Participants rarely discussed the concept of evidence or questioned what the terms “evidence” or “effectiveness” meant when initially responding to the question.

Over the course of the deliberative sessions, participants expressed and debated additional viewpoints. Discussions elicited other interpretations of boundaries, including education or mandates for education, guidelines, accountability mechanisms, and penalties or incentives. Similarly, over the course of deliberation, participants discussed the relative importance of different types of evidence and the role evidence plays in decisionmaking.

Below, we summarize the main themes and values that emerged from the public’s response.

The public’s core values of individual freedom and personal choice were tempered in varying degrees by concern for the greater good or perceptions of fairness.

- The value of individual freedom emerged from participants’ consistent focus on the primacy of personal choice and negative reactions to any boundaries on decisionmaking that restrict rather than support choice. Also, participants often explicitly stated that individual freedom of choice was a core value.
- Concern for the welfare of the community at large arose when discussing evidence that unchecked individual freedom might have consequences that would harm others physically or financially. Since protecting the common good usually entailed some constraints on individual freedom, the conflict between these two values often resulted in discussion about tradeoffs. Reducing individual freedom for the good of the community was not done lightly. Concern for the greater good surfaced most clearly when discussing how blocking inappropriate use of antibiotics could prevent the development of antibiotic-resistant superbugs such as methicillin-resistant *Staphylococcus aureus* (MRSA) or how limiting patients’ choice of hospital to favor a lower volume community facility could enhance a local community’s economic well-being.
- Discussions of health care costs often elicited the value of fairness. Participants viewed fairness from a number of perspectives, including what is just in allocating shared resources and what are reasonable restrictions on patients when they are not the primary payer.

Evidence is an important component of high-quality care. Yet, given the perceived limits of applying population-based evidence to individuals, other factors often have more weight in decisions.

- In general, participants viewed evidence positively and stated that they valued it highly in making their own informed health care decisions. Participants often discussed evidence using terms such as “success rates,” “clinical results,” or “test results.” Other comments indicated that participants equated evidence with experience—the doctor’s accumulated experience and clinical judgment, personal lived experiences, or common sense. Participants’ comments indicated that knowing about unequivocal evidence and uncertain evidence is important when making an informed choice.

- When setting boundaries on decisionmaking, compelling evidence of effectiveness was necessary for encouraging better quality care, but not sufficient for constraining choice or the autonomous decisions of patients and physicians. Yet, if evidence clearly showed a treatment to be ineffective, participants were generally comfortable with setting some restrictions. In comparison, participants could not justify limiting care when the research results are mixed or the evidence itself is unclear.
- Two beliefs emerged that can act to diminish support for the role of evidence in decisionmaking. First was the view that evidence of what works for most people may not apply to each patient, as “everyone is different.” Many comments reflected participants’ perspective that evidence could be discounted if it was seen as “not applicable to me” or not applicable to the unique circumstances of specific patients in specific situations. Second, participants viewed evidence as imperfect: changing over time, often based on studies excluding specific age or ethnic subpopulations, and lacking clarity.
- Other considerations also competed with using evidence in making health decisions. Patients’ personal preferences or doctors’ clinical judgment could supersede evidence. Other features of health care—such as being treated with respect by providers, personal convenience, or concern about out-of-pocket cost—were also instrumental in determining participants’ views. Often, these other factors became more important when participants did not see the relevance of the evidence to the situation.

Evidence of physical or economic harm to individuals or the community led to increased acceptance of some limits on decisionmaking.

- When presented with the deliberative question, many participants’ initial responses showed that they perceived boundaries as compulsory rules and regulations that disallowed exceptions, interfered with the doctor-patient relationship, and limited choice. Participants expressed concerns that boundaries create logistical and practical challenges. Participants also described boundary-setting as a slippery slope, making it easier for future, inappropriate limit-setting.
- Although many participants focused on how to preserve choice and enhance the doctor-patient relationship, the majority of participants eventually concluded that some boundaries would be important or necessary to address problems in the health care system. Descriptions of harm included physical harm (e.g., pain, increased risk of future injury or illness, or death), emotional or psychological harm (e.g., anxiety about outcomes patients can expect), and economic harm (e.g., loss of community jobs, high out-of-pocket expenses for health care). Often, evidence of any harm had a greater influence on increasing acceptance of boundaries than evidence of effectiveness had. In addition, the public perceived outcomes such as death or job loss for individuals in the local community to hold substantial weight and to be more important than inconvenience to a few individuals.
- Evidence of physical harm was the most persuasive factor in accepting boundaries. In most discussions, the preferred way to protect others from harm consisted of guidelines and oversight by medical authorities. In other instances, participants cited and supported rules to prevent adverse effects on the public’s health, such as those now requiring people with tuberculosis to receive treatment.

- Evidence of economic harm was also a persuasive factor in restricting choice. For example, many participants stated that the economic impact and loss of access to care for the community that could result from closing a local hospital, even if it were low performing, outweighed clinical quality for those few who needed specialized surgery. Likewise, participants nearly unanimously supported the need for limits to prevent people from taking advantage of the system and overusing their “fair” share of resources; this was an issue when individual choices increased what others had to pay for health care.

Assessments of risk of physical and economic harm often influenced attitudes about whether society should establish a boundary on decisionmaking: the greater the risk, the more support for the boundary.

- Participants’ perceptions of risk of harm varied, as did the level of comfort with risk-taking. For example, in examining the differences in rates of complications between the low-volume and high-volume hospitals, some participants perceived the level of risk at the low-volume hospital as substantially higher than the risk at the high-volume hospital, while others did not perceive much difference.
- These relative assessments of risk sometimes influenced attitudes about whether society should establish a boundary: the greater the perceived danger, the more support for the boundary.

Although the public believed doctors have the responsibility for knowing and discussing the evidence, they also believed that patients have the responsibility to educate themselves and ask questions of their doctors.

- Participants spoke of doctors’ responsibility to educate themselves about evidence and often identified the doctor as responsible for discussing evidence of benefits and harms with patients so that patients can make informed decisions.
- Most participants believed that patients were responsible for making informed health care choices, asking questions of their doctors, and maintaining a healthy lifestyle. Some strongly supported this perspective from the outset, while others noted that group deliberation changed their views supporting greater patient responsibility.

Doctors—given their understanding of the evidence and the individual patient—should have the authority to determine whether to comply or depart from the evidence in any particular situation. However, the system should hold doctors accountable for their actions to make sure patients receive high-quality health care.

- Participants wanted clinicians to be aware of and generally follow evidence-based guidelines from medical professional associations. Nevertheless, participants believed that clinicians, as experts with specialized education, should be allowed to depart from the guidelines or evidence when needed for individual situations.
- Initially opposed to restricting clinicians’ autonomy, participants often called for increased accountability when faced with evidence that doctors may not always deliver the highest quality care.

Concerns about financial motivations of health care systems, providers, and insurers left many skeptical about whether those setting boundaries or limits in health care would prioritize either evidence of medical effectiveness or quality of care over financial gain.

- Many participants expressed concern that the primary motivation in establishing limits was cost containment rather than ensuring access or quality. Many comments indicated the belief that better care is more expensive and boundaries aimed at cost containment limit access to that better care. Participants were quick to note that costs already constrain patients' choices of and access to certain services.
- Similarly, some participants supported incentives and penalties that could encourage people to adopt healthier lifestyles (e.g., insurance discounts for attending smoking cessation programs) or encourage doctors to provide higher quality care (e.g., professional awards). However, incentivizing physician behavior with financial rewards was more problematic, as participants feared that those incentives might compromise clinicians' integrity by prioritizing financial gain over the patient's health.

The public's trust in entities setting boundaries was influenced by perceptions of expertise, motivation, and whether boundary setting is an appropriate role.

- Overall, participants trusted independent medical associations more than insurers, employers, or government. Participants perceived medical associations as independent, with no financial stake in health care practices or decisions, and as having the needed medical expertise.
- Participants had negative or divided perceptions of other entities based on their perception that such entities lacked medical expertise and/or had questionable motivations. Almost all participants knew that insurers limit care and accepted that as a component of the insurers' role.
- Participants debated whether other payers, such as employers or the government, have the right to set some boundaries. Participants who perceived that these other payers have a legitimate financial or ethical stake in health care tended to accept that these entities could set boundaries. Numerous participants, who had been unaware of the government's large role in paying for health care, became more sympathetic to the idea of government involvement in health care cost containment. Similarly, participants who had been unaware of the risks to society from the overuse of antibiotics tended to become more willing to accept limits on care that promote good antibiotic stewardship.

Throughout deliberation, participants called for more education about evidence and more transparency around health care costs to help inform decisionmaking; some participants even called for government mandates requiring transparent evidence-based information about health care costs, hospital quality, or treatment effectiveness.

- Participants highly supported education and information about health and health care, as most expressed the belief that education and information help people make the best decisions.
- Participants also believed that education about high-quality care is a better approach than restrictive boundaries, especially as education maintains individual freedom and personal choice.

- However, participants held that if education alone is not effective in changing harmful medical practice, then more direct steps for monitoring clinical decisions may be warranted.
- Participants believed that patient access to information about evidence is limited, and a more aggressive effort to bring relevant information to the general public should be a priority. Participants said that the case studies developed for the deliberative discussions would be useful to share with the public: information on provider quality and cost from the hospital quality case study and information on the overuse of antibiotics and MRSA from the URI case study. Participants also wanted general information on treatments and interventions to help improve their decisionmaking.
- Participants noted the difficulties in determining the costs of health care and said that more transparency of health care costs would benefit the public.
- Even though participants generally perceived government interventions that would restrict choice negatively, they typically viewed government mandates requiring transparency, information about costs, and providing evidence-based information about hospital quality and treatment effectiveness positively.

In sum, deliberation required people to consider a variety of tensions and factors in a complex issue, resulting in informed public input that is in depth, nuanced, and actionable. Deliberation allowed participants to explore their own views in more detail, to witness how information and context could influence their perspective and that of others, and to observe how discussion and debate could influence their thinking on the question at hand. As new information or case studies were introduced to the deliberations, answering the overarching question required greater attention to competing priorities. The discussions became more nuanced, with participants exploring the tradeoffs associated with complex individual and societal factors. Although deliberation did not address all misperceptions about evidence or the health care system, numerous participants commented, at the close of their sessions, that they had a deeper understanding of the issues and problems, as well as a better appreciation of a variety of factors relevant to health care.

Effectiveness of Public Deliberation

The randomized design of the Deliberative Methods Demonstration allowed us to assess the impact of deliberation on participants and identify differences by deliberative method and participant characteristics by examining:

- **Changes in participants' knowledge of evidence and comparative effectiveness research.** The knowledge outcome captures the information gained based on questions that were linked to the background educational materials provided to all participants, including the control group. Although participants likely gained additional knowledge from presentations or discussion in deliberative sessions, we measured only the information from the educational materials, which was the most conservative test of increasing knowledge.
- **Shifts in participants' attitudes about the use of evidence in decisionmaking.** Change in attitudes is often measured as an intermediate outcome of effective deliberation. A core assumption of deliberation is that information and discussion may alter the views of

participants as they come to a more informed judgment on the topic. These shifts in attitudes do not have to be for or against a decision; rather, a shift may reflect greater acceptance or greater doubt about one's convictions.¹ Although we used attitude change as a measure of effectiveness, we had no hypotheses for the direction of attitude change. Further, we had no expectation that deliberation would produce group consensus around these attitudes. We assessed attitudes regarding the use of medical evidence in decisionmaking, including questions specific to the hospital quality and URI case studies, and questions on consideration of costs in decisionmaking.

- **Participants' self-reports of the impact the deliberative experience had on them, as well as their assessment of the quality of discourse and implementation.** Impact of deliberative experience included whether participants thought the process affected their views and if participants thought the process was worthwhile. The quality of discourse and implementation included participants' perceptions of the level of participation by all group members, the level of respect for other group members' views, the degree to which participants constructively deliberated the issues, and how well the deliberative methods were implemented.

We assessed these outcomes using two surveys. First, we administered an online survey on knowledge and attitudes to deliberation and control group participants twice, once before educational materials were sent and again within 2 weeks following the conclusion of the deliberative methods. We achieved an 80-percent response rate on the post-survey, using the denominator of all participants recruited (n = 1,774). We summarized knowledge scores as a percent of correct answers. After completing a factor analysis using the attitude items, our final attitude measures included six factors and eight single items.

Second, we administered a survey on deliberation quality and experience one time to participants following their participation, either in person or online depending on the deliberative method. Of the 961 participants who took part in deliberation, 878 participants completed the survey, a response rate of 91 percent. After completing a factor analysis of this survey, our final outcome measures included six factors and two single items.

Below, we summarize findings for five research questions addressing the effectiveness of deliberation and summarize per-group implementation costs for each deliberative method. The unit of analysis for Research Questions 1–4 is the individual participant and for Research Question 5 is the deliberative group.

Public Deliberation Compared With Educational Materials Alone

Question 1: Is public deliberation more or less effective than educational materials alone at changing knowledge about the deliberative topic, and is there a concomitant shift in attitudes?

Participating in deliberation increased participants' knowledge of evidence and comparative effectiveness research.

- Deliberation (for members of all groups combined) increased participants' knowledge of medical issues and concepts related to health care in the United States, the use of medical evidence, and comparative effectiveness research as compared to the control group.

In sum, the increase in knowledge in the deliberative versus control groups represents a clear effect of deliberation on information gained and retained above the use of educational materials alone.

Participating in deliberation shifted participants' attitudes regarding the role of evidence in decisionmaking but did not shift views regarding the relative importance of evidence and personal preferences.

- Deliberation (for members of all groups combined) shifted participants' attitudes related to the importance of medical evidence at a statistically significant level, specifically increasing agreement with:
 - The factor *importance of knowing about medical evidence when making health care treatment decisions*
 - The item *medical research versus doctor's knowledge about patient as most important in medical treatment decisionmaking*
- A shift did not occur in the factor *doctors and patients should consider evidence over preferences when making treatment decisions*.

In sum, deliberation was associated with a shift from agreement to stronger agreement concerning the role of evidence in decisionmaking. When directly proposed against the role of preferences, participants supported the role for evidence, but deliberation did not change views about the relative importance of evidence versus preferences.

When comparing each deliberative method with the control group, all four deliberative methods showed significant change on at least one knowledge or attitude measure.

- Compared with the control group, the CP and BCD methods increased participants' knowledge about evidence and comparative effectiveness research at a statistically significant level. The CD and ODP methods increased participants' knowledge as well, but not at the level of statistical significance.
- Compared with the control group, each of the four deliberative methods shifted participants' attitudes for at least one measure related to the importance of medical evidence at a statistically significant level. For the CP, CD, and ODP methods, shifts showed increasing agreement with the factor *importance of knowing about medical evidence when making health care treatment decisions*. For the BCD and CD methods, shifts showed increasing agreement with the item *medical research versus doctor's knowledge about patient as most important in medical treatment decisionmaking*. For the CP method, shifts also showed increasing agreement with the factor *doctors and patients should consider evidence over preferences when making treatment decisions*.
- Compared with the control group, the CP method shifted participants' attitudes related to considering costs in making treatment decisions at a statistically significant level. Shifts showed increasing agreement with the factor *doctors and patients should consider cost evidence when making decisions*. This factor was evaluated for all methods, as all participants received information on health care costs as context for the discussion. However, the CP method had more time allotted for learning about and discussing issues related to health care costs.

- Attitudes regarding use of medical evidence to restrict antibiotic use reflected a similar impact of deliberation for CP and CD—the two methods that discussed this case study—when each was compared with control. Participants in both methods shifted to more agreement at a statistically significant level on the item *government should limit when doctors can prescribe antibiotics*.

In sum, these findings suggest that all of the deliberative methods can be judged effective compared with a control that used reading materials only on the basis of change on at least one knowledge or attitude measure at the level of statistical significance. However, these statistical tests of individual methods versus control do not allow us to draw conclusions about the relative effectiveness of methods.

Shifts did not occur in three items related to the hospital quality case study, which was used in all methods.

- There were no shifts at a significant level in attitudes related to the material in the hospital quality case study, which all the groups deliberated. This result held true when comparing participants in all deliberative methods compared with the control group, as well as when comparing participants in each method with the control group.

The lack of significant findings may be due to the specific content and complexity of the hospital quality case study. This case study juxtaposed concerns about having access to a “better” high-volume hospital versus the potential impact on the town of having a local low-volume hospital lose business and perhaps close because of reduced patient census. Further, unlike the other case studies, community concerns undermined rather than supported the primacy of evidence.

Overall Quality of Deliberative Sessions

Question 2: What was the overall quality of deliberative discourse and participant experience among the four methods?

Participants reported that they placed a high value on taking part in deliberation and that the experience affected their opinions.

- Participants across all methods placed high value on taking part in deliberation. High ratings of the factor *perceived value of the event* showed that participants valued their participation and included their indication that they would like to participate in activities like this in the future.
- Ratings for the factor *effect of deliberation on participants* reflected participants’ perceptions that the experience had an impact on their opinions on the deliberative topic.

Participants rated the quality of deliberation as high in terms of both the quality of deliberative discourse and the implementation process.

- Participants across all methods rated the quality of communication and discourse highly. Participants reported agreement with the factor measuring the extent that the participants in the groups showed respect for the opinions of others. Participants also reported agreement with the item that people gave reasons to support their opinions. Of note, participants’ ratings for the factor *equal participation in the discussion* were relatively low compared with other measures of discourse quality; participants reported that some people in the group spoke more than others. Despite the fact that participants did not

judge participation to be equal, it did not appear to affect their satisfaction with other aspects of the experience.

- Participants across all methods rated the implementation process highly. Ratings for the factor assessing the quality of the implementation process were overall high, including that the event was well organized, that the information presented was clear and easy to understand, and that the purpose of the event was clear. Ratings for the factor assessing facilitator neutrality were fairly high.

In sum, participants' positive reports of the quality of the deliberative discourse and implementation process indicate that the methods were successful in achieving the core design elements of deliberative methods that were identified in the literature as promoting successful deliberation. Further, positive ratings for the value and effect of deliberation show that participants felt that their input would be used in a meaningful way and that the experience affected them on a personal level.

Comparisons of Specific Deliberative Methods

Question 3: Are specific deliberative methods more effective than others?

Intensity—as measured by contrasting the CP and BCD methods—did not increase knowledge but shifted attitudes at a statistically significant level.

- The higher intensity method (CP) did not increase participants' knowledge of evidence and comparative effectiveness research more than the lower intensity method (BCD).
- Intensity shifted participants' attitudes related to the importance of medical evidence on one factor, *importance of knowing about medical evidence when making health care treatment decisions*, at a statistically significant level. However, intensity did not significantly affect the factor *medical research versus doctor's knowledge about patient as most important in medical treatment decisionmaking* or the item *doctors and patients should consider evidence over preferences when making treatment decisions*.
- The higher intensity method (CP) shifted participants' attitudes related to considering health care costs, specifically increasing agreement with the factor *doctors and patients should consider cost evidence when making decisions*, more than the lower intensity method (BCD).

Intensity—as measured by contrasting the CP and BCD methods—had an effect at a statistically significant level on participants' self-reports of the perceived value of the event, the quality of deliberative discourse, and the implementation process.

- Although participants in both methods placed value on taking part in deliberation, participants in the higher intensity method (CP) reported that the experience had a greater impact on them than participants in the lower intensity method (BCD) did. This difference was at a statistically significant level.
- Participants in both methods rated the quality of deliberative discourse and implementation highly, but differed at a statistically significant level for three outcomes:

- Participants in the lower intensity method (BCD) reported more agreement with the two factors measuring equal participation and facilitator neutrality than participants in the higher intensity method (CP) did.
- Participants in the higher intensity method (CP) reported higher ratings of the quality of the implementation process than participants in the lower intensity method (BCD) did.

In sum, intensity of deliberation, as measured by CP and BCD, has marked impacts on shifts in attitudes and resulted in more positive reactions to the impact of deliberation as reported by participants.

Mode—as measured by contrasting the CD and ODP methods— did not change knowledge or attitude at a statistically significant level.

- Our comparison of an in-person (CD) versus online (ODP) method that required a similar total time commitment from participants did not show a statistically significant effect on any of the knowledge or attitude outcomes.

Mode—as measured by contrasting the CD and ODP methods—had an impact on perceptions of the quality of discourse and impact of the deliberative experience.

- Participants in CD reported significantly higher scores than ODP participants for five out of the eight measures of deliberative experience. For the quality of communication and discourse, CD reported higher scores for the factor *respect for the opinions of others* and the item *reasoned justification of ideas*. For the implementation process, CD reported higher scores for the factor *implementation quality*. For participant reports on the impact of the deliberative experience, CD reported higher scores for the factors *effect of deliberation on participants* and *perceived value of the event*.

In sum, remote (online) methods and in-person methods that engage participants for a similar length of time showed similar changes in knowledge and attitude outcomes. However, our comparison showed dramatic differences between the in-person and online methods in deliberative experience, and specifically around perceived value of the event. This result may be due to the particular nature of our online method, in which facilitation was less active. However, remote methods, regardless of facilitation style, may be less likely to inspire the same level of engagement and excitement as in-person methods.

Participants' Personal Characteristics

Question 4: Does the effectiveness of public deliberation vary by participants' personal characteristics?

Deliberation as a method generally affected people from different demographic groups similarly.

- Regardless of race, ethnicity, age, and educational status, participants showed similar increases in knowledge following deliberation.
- The direction and magnitude of the changes in attitude toward using medical evidence in decisionmaking, including mechanisms to support use of high-volume hospitals, were similar across racial, ethnic, age, and educational lines.

In sum, large and consistent differences among groups on knowledge and attitude outcomes would have suggested that deliberation engaged certain demographic groups more or differently than others. In contrast, we observed no differences in changes in knowledge and few differences in changes in attitude outcomes based on demographic group. These findings suggest that deliberation can be equally effective with a wide range of individuals, not just with more educated or privileged members of the public, as has been suggested in the literature.

Participants from historically underrepresented demographic groups may place more value on or perceive greater impact from their participation than others.

- African Americans and Hispanics reported valuing their deliberative experience even higher than others did.
- African Americans and participants with lower educational attainment perceived deliberation as having a greater impact on their opinions than others did.

In sum, these findings further support deliberation as an effective method for getting input from underrepresented populations.

Concordance—the proportion of a group made up of a specific demographic—generally did not affect participant outcomes.

- We found that concordance was not associated with changes in knowledge among our participants from historically underrepresented groups (African-American or Hispanic participants or participants with lower educational attainment).
- Concordance was also not associated with shifts in attitudes about medical evidence, including use of high-volume hospitals, or with the value or effect of deliberation as perceived by participants.
- However, we did find one exception to this result. For African-American participants, concordance (i.e., the proportion of participants in a deliberative group who were also African American) was associated with higher perceived value of deliberation and also with greater attitude change on the factor *people should consider the effect on group premiums when making treatment decisions* (discussed below).

In sum, we found little evidence that group composition (concordance) affects the shifts in knowledge and attitudes that occur in deliberation. Nonetheless, our findings flag the importance of attention to group composition because of selected findings for African-American participants in groups with higher concordance.

Deliberation highlighted or surfaced select content areas in which demographic groups may hold different views.

- All participants moved from disagreement toward neutral on the factor *doctors and patients should consider cost evidence when making treatment decisions* and the item *people should consider the effect on group premiums when making treatment decisions*. However, there were two differences by demographic group:
 - The magnitude of change on both measures was smaller for African-American participants than for other participants at a statistically significant level for both measures. That is, although all participants moderated their views on the appropriateness of considering costs, African Americans were less inclined than

- others to shift this view. (This result controlled for differences in other demographics, including income and education.)
- Elderly participants changed less than others on the single item *people should consider the effect on group premiums when making treatment decisions*.
 - Hispanic participants agreed more than others before deliberation that doctors and patients should consider evidence over preferences when making treatment decisions. Following deliberation, Hispanic participants' views moderated and their scores drew closer to those of non-Hispanic participants, but they continued to show more support for consideration of evidence over preferences.

In sum, because there were few differences, we conclude that they do not reflect a differential impact of deliberation as a method. However, they suggest some interesting differences in views, which contribute to our findings on the appropriate use of medical evidence.

Internal Dynamics of Deliberative Groups

Question 5: Do the group-level effects (i.e., the internal group dynamics) of public deliberation vary by deliberative method?

There was little systematic movement toward consensus in the Community Forum groups, and none of the methods systematically reached consensus on any of the three measures we used to evaluate consensus.

- For all three measures, only about half the groups moved toward consensus following deliberation, which suggests that achieving consensus was a random—and not inevitable—process.

We found no evidence that polarization—the systematic tendency of groups and the individuals who compose them to strengthen their predeliberation opinions—occurred among any of the methods.

- Following deliberation, 45 percent of the 1,216 observations, or opportunities for attitudes to move toward the extremes, demonstrated movement away from the midpoint and toward the extremes. Because this rate, or opportunity score, is close to 50 percent, it implies that movement toward the extremes occurred randomly and is not systematic or inevitable. There was also no evidence that some measures were more susceptible to polarization than others.

In sum, small-group distortions that have been reported for jury-like settings were not evident in the deliberative groups. We did not find any systematic patterns of polarization (movement away from the midpoint toward the extremes) or movement toward consensus. These results may offer an argument for designing deliberative methods with the core design features that were held constant across methods in our study: no shared consensus seeking and well-tested and balanced educational materials.

Implementation Costs Associated With Holding Deliberative Sessions

The main costs of deliberation include those of developing materials, recruiting participants, holding sessions, and analyzing and reporting results. The costs we report here are limited to those directly associated with holding deliberative sessions; we exclude additional research-related costs we incurred and some other costs we judged to be difficult to generalize. The implementation costs we report include:

- Participant costs, such as incentives or reimbursement for childcare or transportation
- Facilities costs, such as site rental, food, and drink
- Equipment technology, such as microphones, projectors, Internet connection, and telephone conference lines
- Supplies, such as pens, paper, flipcharts, easels, and markers

Our per-group implementation costs are specific to our approach, including a composition of 12 participants per group in BCD, CD, and ODP, and 24 per group in CP. Per-group costs were:

BCD, \$4,500

ODP, \$4,900

CD, \$6,900

CP, \$23,500

For BCD, CD, and ODP, the largest area of implementation cost that we tracked was that of equipment and technology, accounting for more than half of costs. In contrast, the greatest area of cost for CP was participant-related costs (i.e., incentives, transportation, and childcare).

An important factor affecting the total costs of a deliberative project—not reflected in the costs reported above—is the number of groups typically held when implementing a particular deliberative approach. For example, for a given project, BCD usually convenes 10-12 groups, whereas CP may convene only 1-2 groups.

Discussion and Implications

We highlight implications for the two aims of the project that are relevant to entities that use evidence in decisionmaking, as well as those interested in using deliberative methods.

Our analysis of the public's input into the overarching deliberative question highlighted several areas for those entities that generate, translate, or use evidence to inform decisionmaking:

- Our findings show the public's capacity to apply evidence and view health care issues from a societal perspective—and under certain circumstances, to prioritize societal needs over personal ones.
- Given that participants have particular concerns about the impact of harms—and are willing to accept constraints on their autonomy to address harm—effectiveness studies should be as attentive to this domain as they are to evidence of benefit.

- Researchers and policymakers’ concerns about the known limitations of research evidence are shared by the public. These concerns have implications for generation of evidence and translation of research findings.
- To members of the public, more than to other stakeholders in health care, the term *evidence* covers not only the findings of research studies, but also clinical judgment, test results, trial and error, and common sense. The public’s use and understanding of the term “evidence” highlights the complexity and inherent challenges in efforts to translate and disseminate evidence.
- Supporting the lay public’s use and application of evidence requires more than translating the results of scientific studies into plain language. It also requires that clinical evidence be put in the context of other factors when presented to support personal health decisions, such as values, immediacy of results, convenience, or trust in one’s practitioner.
- The public skepticism about the motivations of insurers, employers, researchers, and government involvement in health care suggests the importance of transparency when it comes to disclosing financial interests in health care overall, and specifically in the generation and use of evidence of medical effectiveness.

Our analysis expanded the evidence base concerning public deliberation methods:

- Deliberative methods offer a feasible and effective approach for organizations to obtain informed public views on complex topics affecting broader constituencies. We found that deliberation had similar effects on people, no matter what their race, ethnicity, age, or educational attainment.
- Our overall assessment was that each method was effective. However, the CP and CD methods may be appropriate for more complex topics, while the BCD and ODP methods may be appropriate for less complex topics. Planners will likely want to consider which types of outcomes are most important, as well as the investment required to implement the deliberative method.
- Planners will likely want to consider which types of outcomes are most important, as well as the investment required to implement the deliberative method.
- Because all methods were effective to some extent in eliciting core values, shifting knowledge and attitudes, and having an impact on participants, our overall findings indicate that there is no one right way to conduct public deliberation. Planners who are developing or modifying methods to suit their needs and preferences can weigh the types of tradeoffs we identify and use our results to inform their choices.

Conclusion

Many organizations—researchers, health care providers, and public and private-sector purchasers—as well as multistakeholder efforts to improve community health have an interest in capturing the public voice on complex and value-laden health issues. Further, multiple topics raised by participants over the course of the Deliberative Methods Demonstration—the financing, structure, delivery, and oversight of health care services—are important policy issues undergoing transformations in concept and design at the local, State, and national levels. The

Community Forum Deliberative Methods Demonstration found that public deliberation was an effective, feasible, and useful method to capture public input on these topics.

For More Information

For more information on the AHRQ Community Forum, please contact Joanna Siegel, Sc.D., in the Center for Outcomes and Evidence at joanna.siegel@ahrq.hhs.gov.

The Community Forum Deliberative Methods Demonstration was conducted by the American Institutes for Research under AHRQ Contract No. 290-2010-00005. Contact Kristin L. Carman, Ph.D., at KCarman@air.org for further information. Organizations participating under subcontract included the Center for Healthcare Decisions, Sacramento, CA, and the Center for Deliberative Democracy and the Center for the Study of Language and Information at Stanford University.

Reference

1. Carman KL, Heeringa JW, Heil SKR, et al. The Use of Public Deliberation in Eliciting Public Input: Findings From a Literature Review. (Prepared by the American Institutes for Research under Contract No. 290-2010-00005.) AHRQ Publication No. 13-EHC070-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2013. www.effectivehealthcare.ahrq.gov.