Health Policy Roundtable: Producing and Adapting Research Syntheses for Use by Health-System Managers and Public Policymakers

Growing demand for evidence-based information to inform health care policy and management decisions has inspired new methods for synthesizing relevant information and strategies. This roundtable provides a rationale for the science of synthesizing useful knowledge, including leading-edge initiatives from the United States and Canada.

Chair: Carolyn Clancy, M.D., is the Director for the Agency for Health Care Research and Quality.

Panelists: Linda Bilheimer, Ph.D., Associate Director, Office of Analysis and Epidemiology, National Center for Health Statistics within the U.S. Department of Health and Human Services.

Diane Gagnon, Ph.D., is a Senior Program Officer for granting and commissioning at the Canadian Health Services Research Foundation. **Mark Helfand**, M.D., is Director of the Evidence Based Practice Center at the Oregon Health and Science University and the Portland VAMC.

Carolyn Clancy: Introduction

We have all heard about how much health services research has to offer health policy makers and the public. This roundtable will explore how to distill important research information to enhance its accessibility and usefulness for decision makers. That may sound to some researchers like a project that could be easily handed off to a communicator or PR person—but doing this well is challenging and not straightforward. If it were easy, I am confident that the budget for the Agency for Healthcare Research and Quality (AHRQ) would be at least a billion dollars a year, and AcademyHealth's Annual Research Meeting would have to be held in a facility equipped to handle 20,000 people or more.

As a public funding agency, AHRQ regards research synthesis as a critical part of the responsibility that has been given to us by Congress. We are charged with making the benefits of the work we support tangible and comprehensible to all Americans. Translating research findings into actionable knowledge is an essential goal in health services research. I'm delighted that these three panelists have agreed to share their perspectives on how this can be done. Their efforts represent quite different dimensions of our field, yet provide common themes to help all of us enhance the probability that research will be used to improve Americans' health care.

Linda Bilheimer: The Robert Wood Johnson Foundation's Synthesis Project

I would like to describe a project that I conducted over five years at The Robert Wood Johnson Foundation (RWJF). I came to RWJF from the Congressional Budget Office (CBO), where I learned first-hand about the need to develop models for synthesizing research for the policy community.

As a budget analyst, I was continually frustrated by the mounds of information on my desk. I did not have time to review it all because my analyses and estimates were often due the following morning. Even when I didn't have an urgent deadline, I would often have to wade through 10–20 pieces of research—some from journals, some from interest groups—and weigh the strength of the evidence in order to make the best decision on a given topic. Unfortunately, I had neither the time nor the resources to do that on a routine basis. So the Foundation challenged me to do something about this problem, which many health policy analysts face. The Synthesis Project was the result.

My colleague Claudia Williams and I started by conducting research on our prime target audiences for research syntheses, namely federal policy makers in the congressional and executive branches. Not surprisingly, many of the people we interviewed reported feeling exactly as I did as a federal policy analyst: They were besieged by information and found it difficult to use research effectively.

As a next step, we developed a conceptual model for research syntheses with some specific criteria. We thought that such syntheses should both distill evidence and weigh its strength.

Moreover, they should reflect how policy makers think, and thus be structured around policy questions and draw out policy conclusions. Finally, they should be balanced and succinct.

Following this model, we developed two broad categories of products that are closely linked: *research syntheses*, which are targeted toward

technical and analytical audiences, and *policy briefs*, which are aimed at people who are less analytical or simply have little time to look at new research findings.

We aimed for the research syntheses to be about 18–20 pages in length. Our audience research indicated that the analysts for whom these products are intended want to see data in a compressed form that they can use. They want to be able to extract numbers and confidence intervals easily for their own models and analyses.

Thus, we developed a table-driven model for these more technical products. They include a considerable amount of data and information in tabular form, but we avoid detailed methodological discussions in the body of the text. Rather, that information is contained in appendices that review the methodological issues in the literature as well as the methods that we used to develop the synthesis. The syntheses also include a summary table that describes the literature that has been assessed—its conclusions, whether it has been peer reviewed, whether the methodology is clear, significant strengths or weaknesses of the methodology, etc.

The policy briefs, by contrast, are much shorter, much more visual, and easier to skim. They include simple graphics, which, according to our feedback, their audience prefers over large tables with numbers. Readers can quickly find the policy questions, the synthesis findings, and the policy implications. (Reports and briefs from The Synthesis Project are available at policysynthesis.org. Charts and tables from these products can be extracted easily for use in policy analyses or testimony.)

We are proud of these products, but they are not the primary purpose of the project. Our ultimate goal is to build a new field. We want to help researchers and policy makers to understand the importance of research in informing policy decisions, and to gain academic respectability for the synthesis field as a discipline. The feedback that we currently receive from the academic community indicates that gaining academic credit in this area is extremely difficult. Some junior faculty members say that they would love to get involved with synthesis work but must stick to more traditional research activities in order to obtain tenure.

The national advisory group for the Synthesis Project has been critical in our efforts to establish research synthesis as a unique discipline with academic standards. This group of researchers and policy experts has advised us on every step of the process, and continues to guide our efforts today. For example, members work with us on the selection of appropriate topics that could significantly impact policy debates, based on an initial list that comes from our

target audiences. They also share their ideas about potential authors, and key bodies of literature that the synthesis should include.

National advisory group members also form the backbone of our peer-review process. We divide them into peer-review teams, which collaborate directly with authors from the very beginning of each synthesis. Indeed, the initial meeting between the author and the team is perhaps the most important part of the process. About six to eight weeks after authors begin work on a new synthesis, they meet with their peer-review team to review the outline for the work, discuss which literature should be included, and decide how it should be searched and analyzed.

We wrestled with the question of whether to implement a minimum standard of quality for research that should be included in a synthesis. We tried this approach, and, based on our experiences with our second synthesis, decided that it did not work for us. In that case, a significant body of peer-reviewed literature didn't meet our methodological standards but was being widely used in the policy community to drive decisions. We felt we would not be playing our role as communicators to policy professionals if we did not discuss the strengths and weaknesses of the research that they were already using.

We tried several experiments for communicating the results of these syntheses to the federal policy community, including Web conferencing and closed-door seminars with our target audiences. With both approaches, we found that gathering an audience from the same agency works better than mixing agencies. People seem to feel more comfortable when they are grouped with their peers and colleagues. These discussions also include the researcher who developed the synthesis and a policy person from our national advisory group.

We have benefited tremendously from the emerging work on research translation and synthesis that has come out of Canada and the United Kingdom. I found it particularly helpful, for example, when Nick Mays et al. (2005) drew the now widely accepted distinction between syntheses for decision support, in which research is synthesized to affect decisions, and synthesis for knowledge support, where the goal is to improve the base of knowledge on a given topic for the policy community. RWJF's work is in the area of knowledge support.

We have collaborated with research funders in this country and others to find a lot of common ground and are continuing to grapple with many questions, including the following. When, if ever, is it useful to disseminate the findings of a single study as opposed to a research synthesis? What is the role of researchers and academic reviewers in developing and disseminating synthe-

ses? How can we develop a common terminology and find strategies to evaluate this work?

We are also very interested in finding new ways to communicate effectively to our target audiences. There is much to do as we move forward in developing this critical bridge between research and policy. I look forward to building on our models as well as trying new strategies.

I would like to express my great appreciation to The Robert Wood Johnson Foundation for all the support they have given me in this project and, in particular, to David Colby, who has now taken over my role. I would also like to acknowledge the work of Claudia Williams, the consultant to the project, without whom none of this work would be possible.

Diane Gagnon: Knowledge Transfer at The Canadian Health Services Research Foundation

I am glad to have this opportunity to discuss the syntheses for health systems managers and policy makers that we do at The Canadian Health Services Research Foundation. Research synthesis is a key priority for us; it is one of the main mechanisms we can use to make sure that evidence gets into the hands of decision makers.

The Canadian Health Services Research Foundation was established in 1997 to support evidence-based decision making in Canada's health system. We are a not-for-profit organization that is governed by a board of trustees. We have \$120 million (Canadian) in endowed funds, and our operating budget is about \$10 to \$12 million per year.

Our activities focus on both researchers and decision makers (i.e., managers and policy makers). Our research funding and training programs are designed to ensure that health systems decisions are informed by evidence. Also, we teach managers and policy makers how to assess and use research, and make sure their organizations are capable of taking up research. We conduct a lot of knowledge-transfer and exchange activities as well as networking meetings.

Why should we develop syntheses for decision makers? We've seen evidence that research that is "bundled" engenders more confidence among users than does an individual study. For example, a policy analyst can look at the results of two studies assessing the same research question, one of which has a positive result and the other a negative one, and have no idea which is closer to the truth. If 10 studies are bundled, on the other hand, and one is positive while the other nine are negative, he or she would have a much surer sense of the answer.

There is also the decision maker pull for bundled research. For managers and policy professionals, it is easier to glance at a summary or synthesis of research than to have to track down and assess individual studies. It's less time-consuming, and it provides a better sense of the big picture.

At the Canadian Health Services Research Foundation, we are interested in developing syntheses that aid decision making. We've observed in our discussions with policy makers and managers that many of them are interested in learning the answers to questions that go beyond "What works?" (An example of a "What works?" question that we might address in a synthesis is, "Which evidence-based changes can be made in clinical behavior to improve the quality of care most effectively?") Decision makers are additionally interested in exploring questions about the key sources of influence in patients' decisions to seek care, or whether there are ethnic or socioeconomic differences among patient populations that may affect the allocation of resources for care.

Every three years, the Foundation and other national agencies in Canada travel across the country to consult with decision makers. Our goal is to identify the issues that will affect decision makers' work over the next five years. We then ask our experts to translate those issues into research questions, and to help us distinguish between areas where there are gaps in the research such as the need for more primary studies, and those for which adequate research exists but has not yet been synthesized.

At the Foundation, we have funded nine syntheses to date, and two are in process. One of the two most recent projects is on effective team work in health care and the other is on nurse staffing and patient safety. Since we started doing this work in 1998, we have made many changes along the way and we continue to think about how best to address decision makers' needs.

For example, we initially had few methodological guidelines because research synthesis was an emerging field; thus, each research team would approach their synthesis according to their own take on how to do it. Today, however, we have a bit more guidance about how to go about this process. Indeed, with the NHS Service Delivery and Organization R&D Programme, we commissioned three research teams to look into methodological approaches and explore different ways to synthesize the research for managers and policy makers; their findings were published in the July 2005 supplement of the *Journal for Health Services Research and Policy*, which can be viewed free of charge at <http://select.ingentaconnect.com/rsm/13558196/>.

Over time, we have become more flexible about the timing of the syntheses. In 1998, we believed that each synthesis would take about nine

months from start to finish. When I started working on one, however, I quickly learned that it could be much more time-intensive. It also became apparent that some research questions could be addressed in less time. One issue that arose with the longer syntheses is that they had the potential to lose their relevance by the time they were completed. For example, in one case, a final document was not disseminated until two and half years after the initial question was posed, and we became concerned that decision makers had, by that point, shifted their attention to other policy problems. To prevent this from occurring, we now aim to keep the duration of these projects in the range of 6–18 months.

When we began this work, we asked researchers to develop the policy recommendations that accompanied each synthesis. This approach was sometimes problematic because the researchers' conclusions did not necessarily take into account the specific context within which decision makers were operating. For example, the research team for a synthesis I worked on made recommendations across Canada stating that regional health authorities were very useful in the decision making process. But we were talking to an imperial government that didn't have any regional health authorities at that time, so that created a disconnect between what the research advised and how it could be used.

We also asked the researchers who prepared the syntheses to follow a specific format, which included key messages, an executive summary, and the synthesis itself, so that these documents would be easy for decision makers to access quickly and understand. This was very difficult for many authors, who were accustomed to writing for peer-reviewed journals with a more formal academic tone and format. We hired a journalist to help research teams to communicate more effectively to a policy audience, but the process still required quite a bit of back-and-forth.

To address this issue, we decided to split the process of developing the syntheses into two components. We now ask researchers to work on the syntheses first, draw their research conclusions, and undergo peer-review. In a second phase, we ask a group of decision makers to assess the document and, in collaboration with the research team, discuss the policy implications and develop recommendations.

Over the years, we have had several rounds of evaluation of the syntheses with our board of trustees to assess our progress and consider potential improvements. One of the key issues we have recently discussed is the need to adopt a dissemination approach that addresses the diversity of the decision making world. We decided that a one-size-fits-all strategy won't work, so we

are trying to tailor our approach depending on the decision makers' particular position within the health care system, location, or area of focus. Our plan for doing this is to add customized information and context for different situations and settings after the research conclusions and policy recommendations are complete.

We still face many unresolved questions. How can we balance researchers' need for scientific rigor with policy makers' need for timeliness? What sources of evidence should be considered? What are the best methods for evaluating complex bodies of literature and answering messy questions? Who should write the policy management recommendations? What are the best mechanisms for ensuring that the syntheses are being used?

We'd like to convene a couple of workshops to answer some of those questions. The first would be between researchers and funders to further discuss methodological issues and try to reach a consensus on the optimal approach. Down the road, we want to bring together researchers, funders, and decision makers to discuss these issues as well.

Finding the best way to fund and conduct syntheses for health systems managers and policy makers is a continuing struggle for us, but we are slowly getting there.

Mark Helfand: State-Level Evidence-Based Drug Reviews

I would like to discuss my experience conducting systematic reviews to help state policy makers make judgments about drugs, particularly drugs within a class. This effort began in Oregon, which passed legislation in 2001 to promote the use of cost-effective therapies for Medicaid patients.

In August 2001, John Kitzhaber, Oregon's governor at the time, signed into law Senate Bill 819, which established the evidence-based drug review program. In November, he wrote a letter announcing the program to pharmaceutical companies, in which he emphasized that the effort would be overseen by Oregon's Health Resources Commission (HRC), a citizen panel created in the early 1990s as part of the Oregon Health Plan.

The legislation stated that decision makers in Oregon Medicaid must take into account data about the comparative effectiveness and safety of various drugs. The HRC was charged with making recommendations to Medicaid administrators about which drugs in a class were "preferred." On the provider level, however, this was a totally voluntary mechanism. In other words, although physicians were provided with recommendations through this process, they were free to write prescriptions for the drug of their choice.

The governor offered my colleagues and me at the Oregon Evidence-based Practice Center (EPC) the job of doing reviews of four classes of drugs: COX-2 inhibitors/NSAIDs, statins, proton pump inhibitors, and opioids. We took the job, and in doing so accepted some challenging conditions.

Because this effort was new, state officials didn't really know the cost or value of systematic reviews. The compensation they offered was therefore on the low side. We were offered \$35,000 for each of the four reports. (Our retrospective estimate of the costs was between \$100,000 and \$200,000 each.) In addition, we had to operate on a very short time line; the state initially requested that we generate the reports in three months, although they eventually agreed to six.

The state also requested that we select our research questions by drawing on input from ordinary citizens rather than focusing on what we felt was important. To accomplish this, the HRC appointed volunteer subcommittees for each of the four classes of drugs made up of physicians, pharmacists, and patients or patient advocates.

Finally, state officials told us that all of our work would be disseminated widely before we would have a chance to publish it in a journal. They wanted the entire process to be transparent, and therefore everything we wrote on paper would also be posted on the Web.

We accepted these challenges because we thought making evidence-based reviews on drug classes available to clinicians and the public was an important and original idea.

What are systematic evidence-based reviews? They are comprehensive, open-minded reviews of all the evidence on a given topic. The evidence is used to determine a conclusion, and not the other way around. Often, the alternative to a systematic review is a preselected list of studies that support a preformed conclusion, ignoring or trashing those that don't. In other words, a systematic review has value because it aims to remove bias in finding and selecting scientific studies.

Systematic reviews are also intended to be as inclusive as possible of both positive and negative investigations. Studies with disappointing results often don't get published or receive less attention than positive ones. For example, among five trials of a particular atypical antipsychotic drug evaluated as part of the Food and Drug Administration's pre-approval process, four showed that the drug had some significant effect compared to placebo, while the fifth did not demonstrate any improvement among patients taking the medication. The latter study was the only one that did not get published. Our job is to find that fifth study.

Getting the questions right. Systematic reviews should reflect patients' concerns by addressing the health outcomes they care about. Important questions arise from practice and from life. In a very practical sense, we want these reviews to address vital questions about the cost and effectiveness of drugs to support informed decision making. We don't want the questions to be dictated by what we already know is in the literature, because often the populations, dosing regimens, and measures of effect used in studies are dictated by regulatory requirements, not by the information needs of patients, caregivers, and clinicians.

Medical experts also play a role in these reviews, although they may interpret the data and their own experience differently than researchers. Experts understand the clinical logic behind the research. Without talking to them, it is hard to get the questions right.

Working with partners. When embarking on a project like this, I think it's important to identify the best partners for the effort. What type of policy maker, for example, would collaborate well with researchers and other stakeholders? First of all, I think that person should be a leader. He or she should think about how things ought to be or could be in five years instead of trying to catch up on what should have been done years ago.

For instance, in the hospital world, some managers have gotten noticed because they have had the foresight to think about what information systems might be like 5, 10, and 15 years down the road. By contrast, hundreds of others are still trying to put in place systems that the lead institutions had already implemented 10 years ago. You need to have a policymaker who is really thinking ahead.

On the other hand, I would hesitate to work with a policy maker who wanted to have his hands on everything. You cannot commission systematic reviews if you are a control freak. The basic principle of the process is that you must start with questions, not answers; thus, the policy maker must be comfortable with that uncertainty.

In a similar vein, he or she should not be looking to the researcher to carry out an advocacy agenda. There may be a role for advocacy in the world of health services research. However, these reports are not done to prove a point. They are literature reviews that are conducted in order to separate the procedures and treatments that are based on evidence from those that are not. In our situation, unlike many others, the policy makers involved—Governor Kitzhaber and his colleagues—weren't asking us to arrive at any particular conclusion.

It is also important to find a policy maker who refuses to be pushed around. We've now expanded our efforts to about 13 or 14 other states, which each pay for all or some of the reports we produce. The states have a variety of economic and political landscapes, but one thing they have in common is that they don't want to be pushed around. The way medical evidence is framed and filtered can influence clinicians' beliefs and the public's preferences, and is of great importance to industry and to professional societies.

Systematic reviews level the playing field—any advantage a particular manufacturer may have insofar as reaching clinicians and the public is threatened, at least in part, by a source of information that is not partial to any particular viewpoint. We would not have wanted to be working with policy makers who would pin the blame for anything on us or give in to political pressure.

Now I'd like to explore the characteristics of researchers who would be well suited for this kind of work. First of all, they would have to be willing to work under some of the challenging conditions I mentioned earlier: for possibly less money, on a short time frame, and while disseminating transparent results prior to journal publication. To paraphrase a movie, many researchers might say, "You lost me at 'inadequate compensation.'"

It is perhaps even more difficult to find researchers with health services or clinical epidemiology expertise who are comfortable with policymakers and, in particular, clinicians, actually using the outcome of their work. As others have noted, research is in many cases a dialogue among researchers rather than a guide for real-life decisions. A lot of investigators seem genuinely surprised when they are told that their work will be used as the basis for health or policy decisions.

Researchers must also be able to explain clearly and memorably the principles and rationale for systematic reviews. Anybody who works with policy makers must be a supplier of sound bites. In addition, researchers should be able to speak the other language of policy makers which is, more often than not, anecdotes. Thus, rather than saying that reviews will shift the whole way you think about policy, or casting them as something large, governmental, and population-based, investigators need to bring them down to the level of individual cases and examples, as they do with anything else in policy making.

Systematic reviews and decision making. Some researchers are uncomfortable with the prospect that their work will be interpreted in different ways by different people; they want for there to be one answer and one bottom line

(often, their own). For example, one report focused on the triptans, a class of drugs used to treat migraine headaches. The report found that one of the drugs was more likely to provide relief in one hour than some of the others. However, due to gaps in the evidence, it really wasn't clear after six months or a year which drug within the class of triptans was more likely to improve headaches or enhance functioning. In some states, policy makers concluded that all triptans are essentially the same. In others, officials believed that the one that worked better within one hour was the preferred drug.

It's not that either conclusion is wrong or right; rather, they are both reasonable interpretations of an imperfect body of evidence. Systematic reviews define the strengths and limitations of the evidence, and clarify which tests and treatments are based on evidence and which on other grounds, but they don't necessarily tell you what to do when the evidence is limited. That is where factors such as equity, judgment, values, and preferences come into play. These considerations must be brought to bear to make a good decision—one that reflects the decision maker's values while taking into account the strength and direction of evidence.

I did a lot of work for the U.S. Preventive Services Task Force over the years, and I believe that that prepared me well for comparative effectiveness research. In that position, I would write a report summarizing a body of evidence, which the task force would then use as the basis for their recommendations. At times, I was tempted to tell the group that they had reached the wrong conclusion based on my own review and interpretation of the evidence. But it's not that straightforward. There is a dialogue and you must feel comfortable with that. The report lays out the evidence, but decision makers have to apply their judgment. At times, those judgments may not match your own.

While their conclusions may vary, states and other decision makers all need a process for generating explicit, defensible recommendations and linking them to the evidence. We can't enforce how our reports are used. We can't select for others the right mechanisms to implement systematic reviews fairly and effectively. The most we can do is promote an evidence-based approach. Our challenges are to maintain high academic values, to do our work quickly enough to ensure that the information is timely, to handle criticism well, and to avoid any undue influence.

Finally, we need to identify shared goals among all stakeholders in this process while maintaining separate identities. Clinicians' and patients' concerns should be reflected in these reviews. Researchers and decision makers should wrestle with the idea of what is good evidence. And everybody—

patients, caregivers, policymakers, investigators, payers, etc.—should demand better evidence about outcomes that matter. That is our common agenda.

Prepared for publication by Christina E. Folz, consultant to AcademyHealth.

REFERENCES

- Mays, N., C. Pope, and J. Popay. “Systematically Reviewing Qualitative and Quantitative Evidence to Inform Management and Policy-Making in the Health Field.” *Journal of Health Services Research and Policy* 2005; 10 (3, Suppl): 6–20.