Meta-Evaluation: Using Meta-Analysis and Data Visualization to Inform Policy

Anupa Bir and Nikki Freeman

As the evaluation findings from health care demonstrations and innovations multiply, a wealth of information has become available about what innovative approaches work in particular circumstances. Traditionally, evaluations have sought to answer the question "Does this approach work?" Advancing the discussion to "Would it work for me in my setting?" requires analyzing and synthesizing vast amounts of sometimes contradictory data and the methods used to produce them. In addition, understanding the breadth of settings in which an intervention works or does not work requires an extensive, systematic characterization of the intervention. This brief describes meta-evaluation, an approach that has been adapted by RTI International for its first application to large-scale health care demonstration evaluations. The RTI method developed for this application is inspired by meta-analysis, enhanced by a systematic approach to characterizing interventions, and organized for maximal policymaker utility through data visualization tools. Meta-evaluation allows us to estimate the overall effectiveness of and reasons for variation in large multisite demonstration projects and federal initiatives. This approach makes the lessons learned from complex, multifaceted, real-world demonstrations more apparent and useable for policy decisions.

Inspired by Meta-Analysis

Meta-analysis is a technique for creating a quantitative summary of the evidence surrounding intervention effectiveness across studies. Because not all studies measure outcomes in the same way, meta-analysis begins by transforming the summary evidence from a study, the result(s), to a common metric known as an effect size. Pooling results from multiple studies through meta-analysis improves the power to detect a true effect, particularly when the body of evidence is characterized by small or underpowered studies.

Meta-analytic techniques also provide the tools for quantifying the heterogeneity, or dissimilarity, between study results and for identifying the drivers that explain the differences observed in outcomes. To identify the sources of variation, stratification or meta-regression may be used. Stratification involves performing analyses on different subgroups (e.g., children, adults) and comparing their results. Meta-regression uses linear regression to assess the strength and direction of the relationship between the outcome of interest and intervention-level characteristics—for example, the intervention's main components, setting, intensity

of services provided, target population, or contextual factors. Through these analyses, meta-analysis can model determinants of measured effectiveness.

Enhanced by a Systematic Approach to Characterizing Interventions

To examine the relationships between the effect sizes generated by the different evaluations and the features and implementation characteristics of the interventions, it is essential to plan, develop, and implement the systematic collection of information on the features of interest. This involves

- determining which intervention characteristics to collect, usually by a team with content area expertise;
- identifying and obtaining secondary source information, which in the case of meta-evaluation is often program evaluation findings; and
- creating the necessary infrastructure for collecting the data, including the development, implementation, and surveillance of a standardized data collection protocol.

Once collected, the intervention characteristics can be codified and linked to outcomes of interest. By using

the codified data, common features among the promising findings, even among seemingly dissimilar programs, and potential synergies between the features of effective interventions can be identified. This is also pertinent when evaluation results show no effect, which can reflect an ineffective intervention, poor implementation, or inadequate evaluation design. Our approach seeks to increase the information available to elucidate these relationships.

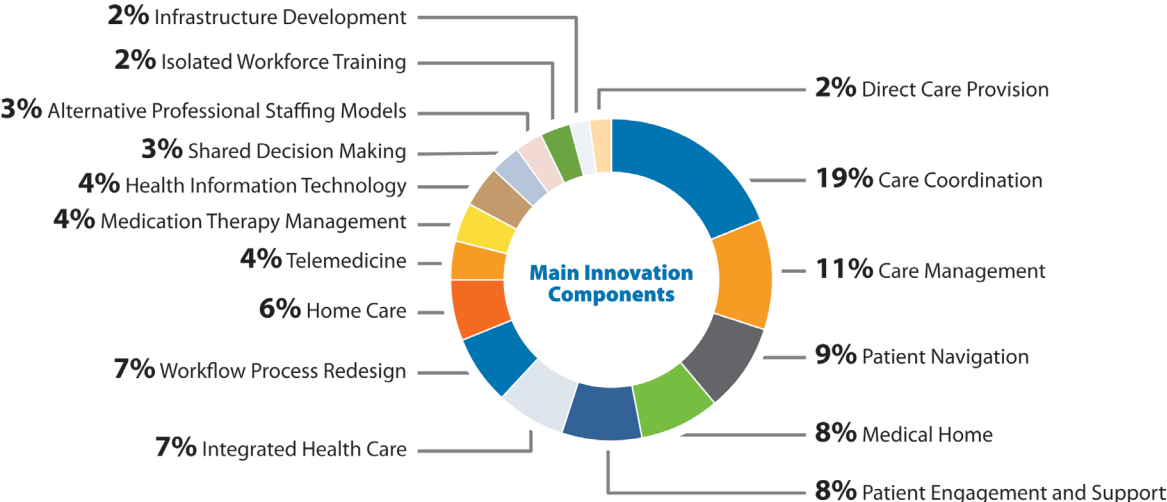## In Practice: Innovative Health Care Payment and Service Delivery Models

The Centers for Medicare & Medicaid Services (CMS) selected RTI to use meta-analytic methods to evaluate, in real time, the performance across 108 Health Care Innovation Awards (HCIA). This program was established as part of the Affordable Care Act in 2010. Congress authorized the Center for Medicare & Medicaid Innovation (CMS Innovation Center) to test innovative health care payment and service delivery models that have the potential to lower Medicare, Medicaid, and Children's Health Insurance Program (CHIP) expenditures while maintaining or improving the quality of beneficiaries' care (42 U.S.C. 1315a). In July 2012, 108 providers, payers, local governments, public-private partnerships, and multi-payer collaboratives received awards ranging from approximately $1 to $30 million for a 3-year period to implement compelling new service delivery and payment models to drive system transformation and deliver better outcomes for Medicare, Medicaid, and CHIP beneficiaries. The initiative was not prescriptive, but rather open-ended, with specific, shared goals of improving outcomes and reducing costs.

Seven organizations—known as frontline evaluators—were selected to evaluate different substantive groupings of HCIA awardees, and RTI was contracted to use the results of those evaluations to estimate the overall effectiveness of innovations and examine why results varied among awardees. To do this, RTI used a combination of rigorous qualitative and quantitative methods to develop a comprehensive understanding of the 108 diverse innovations (see Figure 1) and evidence of their impact on CMS Innovation Center goals, culminating in the development of our meta-evaluation. Qualitative data were obtained from the evaluators' reports and annual awardee surveys completed by evaluation site visit teams. Quantitative data on four CMS-specified core measures (total cost of care, emergency department use, hospital admissions, and hospital readmissions) were abstracted from quarterly and annual evaluators' reports.

Features of each awardee's innovations were systematically collected and coded to identify innovation components and implementation characteristics. We call this "structured coding" to distinguish it from coding that characterizes traditional qualitative research, which seeks to identify themes across awardee reports. Structured coding begins with well-defined, predetermined uniform data elements and an associated coding scheme that a trained team of coders applies to each innovation to result in a standardized data set that characterizes each innovation's target population (e.g., children, adults, elders), setting (e.g., ambulatory, post-acute, long-term care facility), and components (e.g., use of health information technology, provision of care coordination). Additionally, to enhance the set of uniform data elements captured from existing evaluator source material, RTI surveyed the evaluator of each HCIA awardee to systematically examine the degree and intensity of the issues identified in the traditional qualitative coding of evaluator reports. Survey content for this was identified from awardee challenges noted during the qualitative thematic coding using the Consolidated Framework for Implementation Research,[1] a set of widely applicable constructs that have been associated with effective innovation implementation. Together,

Figure 1. Main Innovation Components

- 2% Infrastructure Development
- 2% Isolated Workforce Training
- 3% Alternative Professional Staffing Models
- 3% Shared Decision Making
- 4% Health Information Technology
- 4% Medication Therapy Management
- 4% Telemedicine
- 6% Home Care
- 7% Workflow Process Redesign
- 7% Integrated Health Care
- 2% Direct Care Provision
- 19% Care Coordination
- 11% Care Management
- 9% Patient Navigation
- 8% Medical Home
- 8% Patient Engagement and Support

Main Innovation Components

these methods provided the RTI meta-evaluation team with a rich and comprehensive description of the HCIA awardee innovations, activities, practices, perceptions, and experience implementing and executing their innovations.

In a typical meta-analysis, as much as 25% of the variation in outcome measures is attributable to research methods and procedures.[2] The RTI team provided the awardee evaluators strong guidance to produce methodologically consistent and best-practices results across evaluations. The evaluators used propensity scores—an analytic technique used to mimic the characteristics of randomized controlled trials that produce gold standard evidence in comparative effectiveness research—to create comparison groups from Medicare and Medicaid claims data. Quarterly baseline and post-intervention claims data were used to produce effect sizes for the CMS core measures. Specifically, the effect sizes were generated with the regression difference-in-differences method that isolates the differences between the comparison group and the intervention groups after controlling for changes in measured outcomes that occur naturally over time outside of the intervention.

With the systematically collected data, standard meta-analytic techniques were implemented. As the awardees' efforts are still ongoing, final results are not yet available. When final results are available, the overall performance by the HCIA awardees will again be tested, as will variation in performance by subgroup memberships, such as those providing direct services to patients, those with a behavioral health focus, and those targeting medically fragile populations. RTI will examine variation in estimates and use statistical tools to identify and quantify true variation and range of effects. Meta-regression will be used to identify the innovation characteristics associated with better or worse performance on total cost of care, emergency department use, hospital admissions, and hospital readmissions.

Informing Policy through Data Visualization

Beyond the collection and analysis of data from the HCIA awardees, RTI has developed a dynamic data dashboard to assist policymakers in understanding and using the wealth of HCIA intervention information that has become available through the RTI meta-evaluation. Rather than generating many small graphs and regression tables, the dashboard presents awardee innovation features side-by-side with the corresponding impact outcome data. The dashboard is interactive, and filters allow decisionmakers to focus on features of interest and explore the factors associated with differences in outcomes among interventions.

Figure 2 shows the HCIA dashboard prototype. The HCIA awardees have been filtered to show only the inpatient hospitalization effects, the number of inpatient hospitalizations per 1,000 beneficiaries in a quarter, and characteristics of interventions taking place in the ambulatory setting. The plot on the left shows evaluator-reported estimates of inpatient hospitalization for the awardees. Colors are used to distinguish awardees with statistically significant savings, increased expenditures, and neutral or unknown effects. To the right, awardee characteristics are graphically displayed, representing an array of program features, such as innovation complexity, staffing, and innovation history (i.e., was it a new program or an expansion of an existing program).

Figure 2. HCIA Dashboard Preview
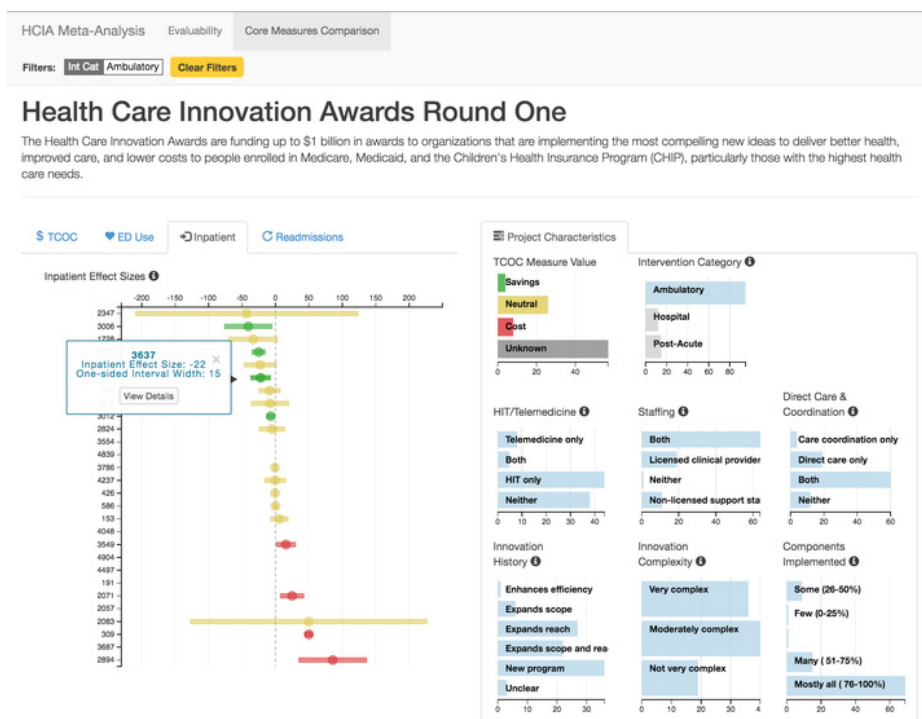
For policymakers and researchers, the dashboard facilitates interaction with the data without requiring programming or statistical expertise. Users can select the characteristics by which to sort and visually compare awardees, empowering them to explore answers to their own questions as they arise. It also provides a way for them to manage data from a vast number of sources. Rather than keep track of volumes of technical reports from a myriad of evaluators and sources, meta-evaluation and the dashboard create a sleek, unified, and responsive tool to examine evaluation findings quickly. Furthermore, it provides decisionmakers with timely information about which characteristics are associated with promising outcomes and may inform decisions about which programs should be scaled up or modified. The goal of the interactive data visualization is to facilitate pattern finding in order to develop models that can be tested.

In addition to creating a single, user-friendly place for combining different types of intervention data, the dashboard's web-based platform enables it to be a dynamic tool that can be updated as new information becomes available. As the meta-evaluation evolves, the dashboard can be extended to include additional variables and modified to examine additional units of analysis, such as additional awardees or more granular analytic units, such as individual practices, when the data allow. It can also be applied to the large-scale demonstration evaluations that encompass hundreds of provider sites, as a support for developing a data-driven, nuanced understanding of interventions and their interaction with implementation contextual factors and outcomes. This dynamism and flexibility is key for policymakers who need to acquire and act on information in real time.

Meta-Evaluation as a Strategic Aid for Policymakers

By combining varied qualitative and quantitative evaluation findings, meta-evaluation allows for the contextualization of quantitative findings and the analysis of individual program components and overall program effectiveness. Although RTI developed meta-evaluation to synthesize HCIA findings, the method has broad applicability. For example, HCIA is a single CMS Innovation Center initiative with many different programs and models being tested for promise in improving quality and reducing cost, but meta-evaluation and the data dashboard could be used to analyze and present findings from multiple initiatives of similar scope and aims. Although various initiatives may seem too dissimilar, meta-evaluation can identify the underlying commonalities between initiatives and link those factors

to program effectiveness, helping policymakers draw upon the lessons learned across initiatives to identify the most promising program components. Of course, conducting meta-evaluation hinges on the availability of evaluation findings reporting similar outcomes across initiatives, and these may come from an amalgam of data sources and evaluators. Sometimes, availability is limited; in these cases, policymakers can harness the power of meta-evaluation to pinpoint inconsistencies or gaps in evaluation questions or outcomes to help policymakers shape future evaluations so they can be the most useful for decision making.

By employing rigorous methods that integrate qualitative and quantitative data with a dynamic data visualization dashboard, RTI's meta-evaluation approach is identifying health care payment and service delivery models that lower the total cost of care while maintaining or improving the quality of beneficiaries' care. For HCIA, the meta-evaluation is identifying the cross-cutting attributes of promising innovations and the drivers of implementation effectiveness in real time. Aggregating, analyzing, and disseminating evaluation findings is a daunting but important task that can provide policymakers a more complete picture of what works and under what circumstances for more fully informed decision making.

References

1. Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science, 4,* 50. doi:10.1186/1748-5908-4-50

2. Lipsey, M. W. (1997). What can you build with thousands of bricks? Musings on the cumulation of knowledge in program evaluation. *New Directions for Evaluation, 1997*(76), 7–23. doi:10.1002/ev.1084

To learn more, please contact:
Anupa Bir, ScD
Senior Director, Health Economics
Center for Advanced Methods Development
+1.781.434.1708
advancedmethods@rti.org

RTI International
3040 E. Cornwallis Road, PO Box 12194
Research Triangle Park, NC 27709-2194 USA