The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Skip all front matter: Jump to Page 1 ▼

## Support RAND

Purchase this document

Browse Reports & Bookstore

Make a charitable contribution

## For More Information

Visit RAND at www.rand.org

Explore the RAND Corporation

View document details

This report is part of the RAND Corporation research report series. RAND reports present research findings and objective analysis that address the challenges facing the public and private sectors. All RAND reports undergo rigorous peer review to ensure high standards for research quality and objectivity.

# CMS Innovation Center Health Care Innovation Awards

## Evaluation Plan

Sandra H. Berry, Thomas W. Concannon, Kristy Gonzalez Morganti,
David I. Auerbach, Megan K. Beckett, Peggy Guey-Chi Chen,
Donna O. Farley, Bing Han, Katherine M. Harris, Spencer S. Jones,
Hangsheng Liu, Susan L. Lovejoy, Terry Marsh, Grant Martsolf,
Christopher Nelson, Edward N. Okeke, Marjorie L. Pearson,
Francesca Pillemer, Melony E. Sorbero, Vivian L. Towe,
Robin M. Weinick

RAND
CORPORATION

# Summary

On November 14, 2011, the Center for Medicare and Medicaid Innovation (CMMI) within the Centers for Medicare & Medicaid Services (CMS) announced the Health Care Innovation Challenge. Through this initiative, CMS planned to award up to $900 million in Health Care Innovation Awards (HCIAs), funded through the Affordable Care Act (ACA), for applicants who proposed compelling new models of service delivery or payment improvements that promise to deliver better health, better health care, and lower costs through improved quality of care for Medicare, Medicaid, and Children's Health Insurance Program (CHIP) enrollees. CMS was also interested in learning how new models would affect subpopulations of beneficiaries (e.g., those eligible for Medicare and Medicaid and complex patients) who have unique characteristics or health care needs that could be related to poor outcomes. In addition, the initiative sought to identify new models of workforce development and deployment, as well as models that can be rapidly deployed and have the promise of sustainability.

This report describes a strategy for evaluating the awardees. It is written for use by the CMS staff who will be engaged in planning the evaluations, by individuals and organizations that will conduct or support evaluation activities, and by awardees who will participate in evaluations by sharing information on HCIA programs and program outcomes. A companion report, *The CMS Innovation Center Health Care Innovation Awards Database Report: Information on Awardees and their Populations* (Morganti et al., 2013), presents detailed information on each of the awardees.

## Goal and Challenges for Evaluation

The goal of the evaluation is to help CMS answer two key questions:

- Which kinds of innovative approaches result in reduced cost while improving or maintaining the standard of care, patient health and quality of life, and satisfaction of the workforce?
- To the extent that a particular approach is promising, what contextual factors need to be in place to make success likely, and what contextual factors might cause serious problems?

There are complex challenges to designing an effective and comprehensive evaluation of the HCIA initiative. Below, we summarize a few of these challenges for evaluation design and implementation. All of these challenges will be addressed in the proposed strategy:

- **Evaluation skills of individual awardees.** Based on past experience with multiproject evaluations, we expect that the awardees funded under HCIA will

vary widely in terms of their capacity to provide information for carrying out the external evaluation. In order to conduct the evaluation, the external evaluator(s) will need to assess awardees' skill levels and establish a priority list for technical assistance.

- **Calculating program costs and cost savings.** A primary focus of program evaluators will be to identify models with the potential to reduce the total cost of care in ways that are financially sustainable for provider organizations. Key to achieving this aim is the ability of evaluators to measure effects of a comprehensive array of innovations on cost and quality outcomes. The ability to calculate program costs and cost savings will be challenging for many awardees.
- **Awardee time burdens and coordination.** Both in this evaluation design contract and the subsequent contract to perform the evaluation, it is essential to balance the need for valid evaluation data against the desire to minimize time and resource burdens on individual awardees. Our experience has demonstrated that it is preferable (whenever possible) for the evaluator to conduct secondary analyses of existing data and to supplement these analyses with awardee interviews or additional materials directly from their projects. This approach enables the evaluator to obtain complete data and eases the evaluation burden on awardees.
- **Cross-cutting versus awardee-specific measures**. A crucial issue for the evaluation strategy is to achieve a balance between the use of measures that cut across awardees' efforts and those specific measures relevant to each innovation model being tested. The correct balance will ensure that measures are general enough to inform cross-awardee comparisons while being relevant to a diverse array of awardee efforts and will also ensure that awardees do not bear an undue measurement burden.

## An Evaluation Strategy for HCIA Awardees

The goal for the evaluation design process is to create standardized approaches for answering key questions that can be customized to similar groups of awardees and that allow for rapid and comparable assessment across awardees. The evaluation plan envisions that data collection and analysis will be carried out on three levels: at the level of the individual awardee, at the level of the awardee grouping, and as a summary evaluation that includes all awardees. The ultimate goal is to identify strategies that can be employed widely to lower cost while improving care.

### Evaluation at the Level of the Individual Awardee

The first step in conducting an evaluation for each awardee will be to develop data at the level of the individual awardee. This may involve collection of program documents and other

materials, clinical data, self-report data from program patients or staff, or observational data (e.g., observations of key program activities being implemented).

*Evaluation Within Awardee Groupings*

In order to conduct evaluations at an operationally manageable level and to allow potential pooling of data for statistical analysis, RAND developed and CMS reviewed and approved groupings of awardees. We proposed a way of grouping awardees based on the larger questions the evaluation needs to answer, as well as on the day-to-day realities of how and in what parts of the care system the awardees are implementing their projects (i.e., approach and setting). We suggested grouping awardee projects across three larger categories:

- **Management of medically fragile patients in the community:** characterized by a focus on high-risk patients who are receiving care primarily in community health care settings with a goal of controlling costs by improving care quality and reducing emergency department (ED) visits and hospital admissions
- **Hospital setting interventions:** characterized by a focus on care of hospitalized patients with a goal of reducing the length of stay, intensity of utilization, duplication of care, and readmission
- **Community interventions:** characterized by a focus on care of beneficiaries in community settings, but focusing on various aspects of how care is delivered, rather than certain categories of patients—although some may also focus on subgroups of patients

While these three types of approaches are designed to improve quality of care and reduce or slow the growth of cost through better care, they will do so in different ways and with different specific end points, and these differences will need to be taken into account in designing an evaluation plan. It will also be important to capture the specific structural features of programs (e.g., health information technology [HIT] improvements, workforce training, payment reform); the processes they include (e.g., care coordination, patient navigation, home visitation, care standardization); the effects on specific clinical outcomes and health-related quality of life; and the specific ways in which they are affecting cost in terms of reduced intensity of care, reduced ED visits, reduced hospitalizations and readmissions, and other factors.

We proposed ten groupings for awardees within these three categories, as shown in Table S.1. Following discussions with CMS about the proposed groups and the assignment of awardees to the groups, RAND worked with CMS to finalize the assignment of awardees to the ten groupings.

**Table S.1. Summary of Awardees Categories and Groupings**

| Category | Groupings |
|---|---|
| Management of medically fragile patients in the community | • Disease/condition-specific targeting (e.g., cardiac, asthma, dementia, diabetes, stroke, cancer, chronic pain, renal/dialysis)<br>• Complex/high-risk patient targeting (e.g., multiple conditions, rural, low income, advanced illness)<br>• Behavioral health patients being treated in community care settings |
| Hospital settings interventions | • Condition-specific targeting (e.g., sepsis, delirium)<br>• Acute care management<br>• Improvement in ICU care, remote ICU monitoring |
| Community interventions | • Community resource planning, prevention and monitoring<br>• Primary care redesign<br>• Pharmacy/medication management<br>• Shared decisionmaking |

NOTE: ICU = intensive care unit.

In addition to the grouping structure for the awardees, there are other characteristics that will be considered in the evaluation design recommendations and for the actual evaluation. These include

- target population characteristics (e.g., age, Medicare status, Medicaid status, and CHIP status)
- geographic region of the country and whether the area is urban or rural
- program size in terms of funding and number of beneficiaries
- workforce characteristics (e.g., type of staff, kinds of training and organization, and how types and levels are staff are deployed to undertake tasks within the system of care).

## *Evaluation Across All Awardees and Groupings*

The value of a summary evaluation is the opportunity for CMS to examine aspects of program implementation, workforce, and context that may influence an intervention's effectiveness. We present several approaches for a summary evaluation of awardees and groupings. These include a meta-analytic approach, pooled data analyses, and a systematic ratings system. These approaches will help to identify intervention strategies that are most effective in reducing costs while improving quality of care. Finally, we present structured approaches for establishing consensus interpretations of awardee and grouping evaluations, as well as for arriving at decisions about which approaches are worth scaling up, which are worth studying further and which should be deferred from current consideration for further investment.

## Evaluation Dimensions, Measures, and Designs

### *Conceptual Framework*

The conceptual framework for the evaluation is shown in Figure S.1. The framework illustrates how key dimensions of the evaluation relate to a primary outcome of interest: the sustainability of an awardee program.

In the leftmost box, we depict the *health status and characteristics* of the target patient population. These characteristics motivate the *design of an innovation program*, which is also influenced by the *legal, regulatory, and fiscal environment*; the *organizational context*; and the *workforce context*. The *implementation effectiveness* of each program is affected by *organizational context* and *workforce training* and can be measured along four dimensions: *program drivers* (i.e., the theory behind the program and intended drivers of change); *intervention components* (e.g., training, technical assistance), *dosage* (i.e., the "amount" of the intervention delivered to patients or the health system), and *fidelity* (i.e., adherence to planned procedures); and the reach of the program. *Program effectiveness* is characterized by the evaluation dimensions of *health*, *cost*, and *quality*. All of these factors affect the *return on investment* (ROI), which, along with *workforce satisfaction*, affects the overall *sustainability* of the program. Each dimension in this framework represents a more complex set of elements. This framework is meant to be flexible so that it can be operationalized and interpreted by stakeholders with varying perspectives, including providers, evaluators, and CMS.

**Figure S.1. Conceptual Framework**
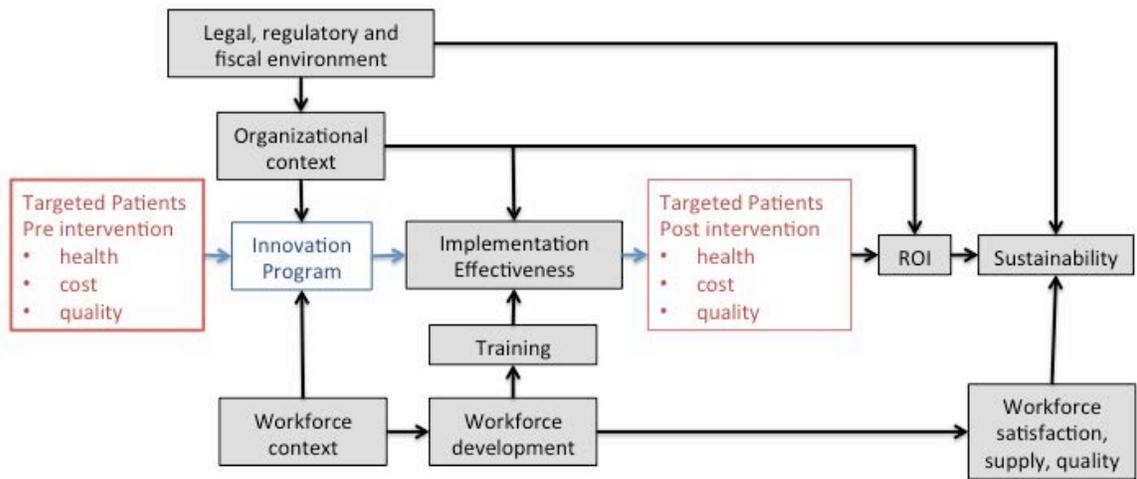


## Key Dimensions Within the Framework

In Table S.2, we outline the key dimensions for the proposed evaluations. Below the table, we briefly define each of the dimensions and its importance for the HCIA project and explain the focus of measurement for the dimension.

**Table S.2. Evaluation Dimensions**

| Category | Dimensions | Subdimensions |
|---|---|---|
| **I. Implementation Effectiveness** | | |
| | A. Program drivers | 1. Theory of change |
| | | 2. Theory of action |
| | B. Intervention | 1. Components of the intervention |
| | | 2. Dosage |
| | | 3. Fidelity |
| | | 4. Self-monitoring |
| | C. Reach | 1. Coverage |
| | | 2. Timeliness of implementation |
| | | 3. Secondary use of tools |
| **II. Program Effectiveness** | | |
| | A. Health | 1. Health outcomes |
| | | 2. HRQoL |
| | B. Costs | 1. Program costs |
| | | 2. Utilization |
| | | 3. Expenditure |
| | C. Quality | 1. Safety |
| | | 2. Clinical effectiveness |
| | | 3. Patient experience |
| | | 4. Timeliness |
| | | 5. Efficiency |
| | | 6. Care coordination |
| | D. Cross-cutting considerations | 1. Equity and disparities |
| | | 2. Subgroup effects |
| | | 3. Spillover effects |
| **III. Workforce Issues** | | |
| | A. Development and training | |
| | B. Deployment | |
| | C. Satisfaction | |
| **IV. Impact on Priority Populations** | | |
| | A. Populations | 1. Medical priority groups |
| | | 2. Nonmedical priority groups |
| | B. Impact | 1. Cost reductions and savings |
| | | 2. Clinical outcomes |
| **V. Context** | | |
| | A. Endogenous factors | 1. Leadership |
| | | 2. Team characteristics |
| | | 3. Organizational characteristics |
| | | 4. Stakeholder engagement |
| | B Exogenous factors | 1. Policy and political environment |

## Implementation Effectiveness

Implementation effectiveness refers to the degree to which an intervention is deployed successfully in real-world settings. Speed to implementation was a key consideration in the selection of HCIA awardees, and a key goal of the HCIA program is to identify innovations that can be rapidly deployed more widely once they have been determined to be effective.

Implementation effectiveness can be measured in terms of program drivers; intervention components, dosage, fidelity, and self-monitoring; and reach. *Program drivers* include the theory of change (i.e., the mechanisms that catalyze or otherwise cause changes in individual and organizational behavior) and the theory of action behind the intervention (i.e., the specific activities used to deliver the innovation). *Intervention components* include the specific activities by which the program seeks to induce better health outcomes at lower cost (e.g., training programs, patient navigators, HIT, new staffing). *Dosage* refers to how much of the innovation a health system or patient gets. *Fidelity* refers to how faithfully the innovation or program was delivered. *Self-monitoring* refers to awardee efforts to collect data on their own program activities and outcomes and the use of these data for quality improvement. *Reach* can be measured through the extent of the intervention's coverage (i.e., geographic reach, target population, number of individuals, organizations, or other units covered), the timeliness of its implementation, and the secondary use of tools that it generates.

## Program Effectiveness

Program effectiveness refers to assessments of an intervention's impact on outcomes of interest, referring to the goals of reducing cost through better care and better health. HCIA awardees are expected to assess cost savings and to document improvements in health outcomes and quality over the three-year term of the award. They are also asked to project the intervention's effectiveness on an annualized basis after the term is finished.

We present three outcome dimensions that are of interest in health care innovation: health, costs, and quality. The health dimension focuses on the impact of the intervention on health outcomes, including mortality, morbidity, and health-related quality of life (HRQoL). The costs dimension focuses on program costs, impact on utilization, and changes in expenditures resulting from the intervention. The quality dimension focuses on improvements in care along several domains of quality: (1) safety, (2) clinical effectiveness, (3) patient experience, (4) timeliness, (5) efficiency, and (6) care coordination. We also discuss considerations that cut across the other dimensions in this section—including equity and health care disparities issues, effects on specific subgroups of interest, and spillover effects.

## Workforce Issues

A critical challenge of delivery system reform is to identify and test new ways to create and support the workforce of the future—a workforce that will deliver and support new care models.

There are three key types of workforce issues to be considered: development and training, deployment, and satisfaction. In terms of development and training, it is important to understand what works best for implementation of the innovation: a training process and other strategies to add new skills to current workers or contracts with outside providers who already have those skills. How workers are deployed and how they interact with patients is also critical to the success or effectiveness of many of the awardees' interventions. Job satisfaction is key to providers' willingness to be part of this workforce, their ability to perform their work effectively, and the smooth functioning of a provider organization.

Key elements of development and training to be measured include the extent to which programs provide training to use existing staff and incorporate new kinds of staff effectively, the level of investment in training required to fill workforce gaps, and the effectiveness and efficiency of various training models. Deployment issues include the extent to which newly formed teams function together and the ways in which workforces are utilized in the innovation. To understand staff satisfaction, it is important to measure the extent to which different kinds and levels of staff are satisfied or dissatisfied with the care they are able to provide and with working conditions in general.

## Impact on Priority Populations

Priority populations may include those with certain medical conditions, such as the chronically ill, pregnant women, persons with behavioral health needs, individuals with disabilities, and people living with HIV. Nonmedical priority populations might include senior citizens, children, low-income families, homeless individuals, immigrants and refugees, rural populations, ethnic/racial minority populations, non–English-speaking individuals, and underserved groups. Evaluating the impact of HCIA interventions on priority populations means understanding the potential impact of the intervention on these populations, including the impact on clinical outcomes and cost.

Two aspects of measuring intervention impact for priority groups are important: (1) the extent to which health outcomes, quality, and costs are different for individual priority groups compared to the health outcomes quality and costs for the intervention population as a whole and (2) whether outcomes, quality, and cost savings would be different for priority groups if the intervention were brought to full scale.

A number of metrics might be used to measure outcomes for priority groups. These include patient characteristics, mortality, morbidity, functional health status, HRQoL, technical quality, rating of providers, rating of provider communication, access to care, care coordination, courtesy and helpfulness of providers, cultural competency, self-management education, and rating of experience with new technologies and processes. In addition, it will be crucial to understand how cost impacts and population size may interact to produce potential savings.

Context

Context refers to the environment in which an innovation occurs and, more specifically, to the factors that can help facilitate or impede an intervention's success. Context includes such endogenous factors as leadership, team functioning, and organizational features and such exogenous factors as the policy and political environment in which an intervention is implemented. Key questions focus on the contextual factors that are needed to support a particular intervention: Were there unique characteristics of the awardee organization, market, approaches, or patient populations that affected the implementation and success of the innovation? Was there a clearly designated champion or leader to oversee implementation?

Key dimensions of context to be measured include endogenous factors (i.e., awardee characteristics, programmatic changes, leadership, team science, organizational issues) and exogenous factors, such as the policy and political environment. The relevant aspects of context will vary across interventions. Because they vary, we propose to assess context in terms of "fit" or "congruence" between two key elements: the demands and requirements of the innovation and the operational realities of the use context.

## Summary Evaluation and Decision Strategy

In addition to the evaluations of individual awardees and awardee groups, we also see a role for summary evaluation strategies that would include other awardee groupings. For instance, a summary evaluation might assess awardees that include Medicare recipients as their primary target group. The primary objective of the summary evaluation is to compare and synthesize findings from evaluations conducted at the awardee and group levels, as well as from pooled analyses. The evaluations will assist in identifying (1) those interventions that can be implemented more broadly, (2) those that need testing in other settings, and (3) those that may be deferred from current consideration for further investment.

The benefits of a summary evaluation have to do with the potential to compare, synthesize, and interpret the variety of evaluations that are conducted on individual innovations and smaller groups of awardees. Comparison and synthesis can provide further insight on innovations that are effective at controlling or reducing costs and those that are effective at maintaining or improving health outcomes and quality of care. A summary evaluation can also provide data on how effective innovations can be scaled up to other populations and under what circumstances; what changes in regulations, reimbursement structure, and other policies may be needed to ensure the sustainability of effective innovations; and how less-effective innovations can be tested further, why their outcomes are lacking, and how their outcomes might be improved.

There are also several challenges associated with conducting a summary evaluation. The first of these has to do with the heterogeneity of awardee activities. Each awardee has proposed and is carrying out multiple, overlapping changes in its health care systems. Second, the awardees target a wide range of populations, and thus care must be exercised in interpreting the potential

for scale-up of successful innovations. Third, awardee innovations and their population impacts will be evaluated in the context of different organizational characteristics (e.g., differences in leadership support, information technology [IT], culture, staffing structure), which may be influential on outcomes. Fourth, and perhaps most challenging, individual awardees and evaluators may measure performance in different ways, which means that comparison and synthesis of measurement will be extremely challenging.

## Strategy

The summary evaluation strategy has to take account of these challenges. Below we suggest key elements of a strategy that will create opportunities for valid comparison and synthesis of individual awardee and group evaluations.

### Coordination of Evaluators

Early coordination of evaluators will be important because it can maximize correspondence and minimize unnecessary variation in the ways that awardee innovations have been assessed, through differences in evaluation questions, metrics, data, or approach. As awardee and group evaluations proceed, coordination will ensure that questions, metrics, data, and approaches are similar enough to produce findings that can be compared and synthesized across the many awardees, awardee group, and interventions. Coordination would begin with consideration of proposed evaluation dimensions. The process would continue with a discussion of the research questions, metrics, data, and approaches for evaluation within each of the awardee groupings.

### Analysis and Interpretation of Findings

The analysis and interpretation approach we propose is composed of three major components, which can be carried out simultaneously.

**Component 1: A Ratings System.** An evaluation ratings system may be developed to compare findings from the many qualitative and quantitative measures in grouping, intervention, and program evaluations. This system could be focused on the five major evaluation dimensions presented earlier: implementation effectiveness, program effectiveness, workforce issues, impact on priority populations, and context. The characteristics are designed to summarize findings across evaluation dimensions, using different types of data.

**Component 2: A Pooled Analysis.** Further assessment of the interventions undertaken by awardees can be obtained via a pooled analysis using data from CMS, states, or other administrative or survey sources. The power of a pooled analysis is to combine observations from multiple awardees to enhance statistical power and isolate the effects of different interventions by controlling for features that vary across interventions. This pooled analysis would likely focus on program effectiveness and the subdimensions of health, costs, and quality. Although it can add further insight into the performance of *individual* awardees, the main strength of a pooled analysis is to shed light on the effectiveness of certain *types* of interventions

and how that effectiveness is influenced by other factors, such as setting, context, or populations involved in the intervention. The strength of the analysis depends on the availability of suitable control populations and standardized and timely data on the individual interventions. The pooled analysis is designed to identify key elements of implementation effectiveness by taking advantage of larger sampler sizes and comprehensive analytic techniques.

**Component 3: A Decision Strategy.** The qualitative and quantitative comparisons and syntheses in Component 1 will address opportunities for cross-awardee learning in each of the five dimensions presented above. The pooled analyses from Component 2 will focus on program effectiveness and its subdimensions of health, costs, and quality, taking into account opportunities for pooling CMS, state, and other administrative data. A structured decision strategy would use data from these first two components to enable systematic consideration of key innovation features and outcomes to develop informed policy. The comparisons and syntheses that arise from pooled analyses have the potential for stronger internal and external validity of findings in the summary evaluation. These pooled analyses can thus be seen as an independent validation of findings from individual awardee, grouping, and Component 1 evaluations.

A summary evaluation may be carried out concurrently with the individual awardee and group evaluations. In order to accomplish this, the evaluators need to be coordinated in their work and have a clear plan for analysis, synthesis, and interpretation of their results.

## Conclusion

The CMMI investment in new care and payment models is of potentially historic importance in terms of how we control health care costs while improving quality and outcomes. The evaluation of these awards will inform decisions about expanding the duration and scope of the models being tested. Despite the challenges, the evaluation and decision process must be of the highest technical quality, as well as transparent and well communicated. Thus, evaluators will have a critical role in the effort to reduce costs while maintaining quality in the delivery of health care. The strategy proposed in this report is put forward with these challenges in mind.