

**RAPID EVALUATION
APPROACHES FOR COMPLEX
INITIATIVES**

WHITE PAPER

March 31, 2014

Margaret B. Hargreaves, Ph.D., M.P.P.



MATHEMATICA
Policy Research

This report was prepared by Mathematica Policy Research, under contract to the Assistant Secretary for Planning and Evaluation. The findings and conclusions of this report are those of the author and do not necessarily represent the views of ASPE or HHS.

Contract Number:
Building Analytic Capacity for
Policy Analysis and Decision Making
HHSP23320095642WC/ HHSP23337039T

Mathematica Reference Number:
40146.510

Submitted to:
U.S. Department of Health and Human
Services,
Assistant Secretary for Planning and
Evaluation
200 Independence Ave, SW
Washington, DC 20201
Project Officer: Amanda Cash

Submitted by:
Mathematica Policy Research
955 Massachusetts Avenue
Suite 801
Cambridge, MA 02139
Telephone: (617) 491-7900
Facsimile: (617) 491-8044
Project Director: Marsha Gold

**RAPID EVALUATION
APPROACHES FOR COMPLEX
INITIATIVES**

WHITE PAPER

March 31, 2014

Margaret B. Hargreaves, Ph.D., M.P.P.

MATHEMATICA
Policy Research

CONTENTS

OVERVIEW.....	v
I. INTRODUCTION.....	1
II. REVIEW OF RAPID EVALUATION APPROACHES.....	2
A. Performance Measurement and Quality Improvement	2
B. Rapid Cycle Evaluation.....	3
C. Complex Systems Evaluations	4
III. RAPID EVALUATION COMPARATIVE FRAMEWORK.....	6
A. The Context	8
B. The Intervention.....	9
C. The Evaluation.....	10
IV. PROCESS CHANGE—QUALITY IMPROVEMENT	11
V. ORGANIZATIONAL CHANGE—RAPID CYCLE EVALUATION	14
VI. COMPLEX SYSTEMS CHANGE—SYSTEMS CHANGE EVALUATION	16
VII. CONCLUSIONS.....	20
BIBLIOGRAPHY	21

TABLES

1	Rapid Evaluation Dynamics Comparative Framework.....	8
2	Rapid Evaluation: Simple Process Change.....	13
3	Rapid Evaluation: Complicated Organizational Change.....	16
4	Rapid Evaluation: Complex Systems Change.....	19

FIGURES

1	Continuous Quality Improvement.....	11
2	Rapid Cycle Evaluation.....	14
3	Nested, Systems-based Evaluation.....	17

OVERVIEW

This paper addresses the challenges of conducting rapid evaluations in widely varying circumstances, from small-scale process improvement projects to complex, system transformation initiatives. Rapid approaches designed to evaluate projects at lower levels of complexity do not take into account the inter-organizational aspects of more complex initiatives, especially those designed to build capacity and integrate activities across organizations, sectors, and levels. Providing a framework that recognizes key differences in the scope and complexity of interventions helps to advance implementation science beyond a program-centric focus on process and organizational improvements to encompass a whole systems approach.

This paper presents a comparative framework of rapid evaluation methods for projects of three levels of complexity: quality improvement methods for simple process improvement projects; rapid cycle evaluations for complicated organizational change programs, and systems-based rapid feedback methods for large-scale systemic or population change initiatives. The paper also provides an example of each type of rapid evaluation and ends with a discussion of rapid evaluation principles appropriate for any level of complexity. The comparative framework is designed as a heuristic tool rather than as a prescriptive how-to manual for assigning rapid evaluation methods to different projects. No one best rapid evaluation method works in all circumstances; the right approach addresses the goals of the evaluation and captures the complexities of the intervention and its environment.

Different rapid evaluation methods are appropriate for different circumstances. Quality improvement and performance measurement methods are appropriate for process improvement projects. Rapid program evaluation methods are appropriate for organizational change programs. Developmental evaluation and other systemic change evaluation methods are best for larger-scale systems change initiatives. However, these methods are not mutually exclusive. They may be more effective when nested, just as simple checklists are used to reduce preventable surgical errors within larger hospital safety campaigns that are funded through national payment reforms that reward system-wide shifts in health care costs and quality. Evaluating an intervention from process, organization and systemic perspectives allows us to implement change more effectively from multiple vantage points.

Regardless of the level of complexity, rapid evaluation methods should maintain a balance between short-term results and long-term outcomes so that there is an alignment of task control, management control, and strategic control. They should also be part of an interactive and adaptive management process in which internal operational results and external environmental feedback are used together in an iterative process to test and improve the initiative's overall strategy.

I. INTRODUCTION

The U.S. Department of Health and Human Services (HHS) embarked on an effort to build internal evaluation capacity throughout the agency and its multiple components. HHS is among many federal agencies aiming to build stronger internal capacity, in large part due to internal HHS efforts and as a result of the federal Office of Management and Budget's (OMB) memorandum in 2012 encouraging the use of evidence and evaluation in budget, management, and policy decisions to make government work more effectively. HHS has an internal evaluation working group that has worked to outline the practical value of evaluating HHS programs as well as common challenges encountered when considering the fit of an evaluation approach with different types of programs. Multiple evaluation approaches are used across HHS, often in combination, to address complex questions about program implementation, whether a program, policy, or initiative is operating as planned and achieving its intended goals, and why or why not.

Across many HHS agencies, the first step in an evaluation is determining the type of evaluation that can reasonably be conducted. Through its work, the HHS evaluation work group identified a need for new methods to conduct more rapid assessments and evaluations of programs as they are being implemented, rather than waiting for years after programs have ended for evaluation results. The federal government is also looking for new ways to work more efficiently and effectively to evaluate programs within an evidence-based framework. At the same time, policymakers, health care providers and public health practitioners are employing multifaceted interventions targeting large-scale organization and systems change at multiple levels in health care, behavioral health, public health and human services. The motivation for this paper is to articulate appropriate rapid evaluation methods for such complex and multilayered initiatives.

To prepare this paper, Mathematica Policy Research gathered information from colleagues about current efforts to conduct rapid cycle evaluations of large-scale initiatives, collected and reviewed relevant literature on rapid evaluation methods and related approaches, and attended rapid evaluation methods sessions at the June 2013 AcademyHealth Meeting to learn more about the methods and findings of current rapid evaluation projects. To analyze the information, Mathematica compared the rapid evaluation methods used in projects of differing complexity and their key attributes, and identified evaluation projects exemplifying different approaches.

This paper addresses the challenges of conducting rapid evaluations in widely varying circumstances, from small-scale process improvement projects to complex, system transformation initiatives. Rapid approaches designed to evaluate projects at lower levels of complexity do not take into account the inter-organizational aspects of more complex initiatives, especially those designed to build capacity and integrate activities across organizations, sectors, and levels. Providing a framework that recognizes key differences in the scope and complexity of interventions helps to advance implementation science beyond a program-centric focus on process and organizational improvements to encompass a whole systems approach (Perla 2013).

The paper is organized as follows. Section II provides a review of various rapid evaluation approaches that were developed for different kinds of initiatives. Next, Section III presents a comparative framework of rapid evaluation methods for projects at three levels of complexity: quality improvement methods for simple process improvement projects, rapid cycle evaluations

for complicated organizational change programs, and systems-based rapid feedback methods for large-scale systemic change or population health initiatives. Then, Sections IV, V, and VI provide examples of each type of rapid evaluation. In Section VII, the paper ends with a discussion of the value of rapid evaluation principles that are appropriate at any level of complexity.

II. REVIEW OF RAPID EVALUATION APPROACHES

There are numerous rapid methods for evaluating program, system, and organizational change. Each approach was developed to address particular problems or improve certain kinds of practices. However, they have often been adopted for universal use without consideration of the circumstances for which they were originally created. This section provides a brief overview of a sample of the better known rapid evaluation approaches that are often used in HHS studies, focusing on their purposes, methods, and uses. This is not a comprehensive description of every method; the overview is intended to provide a context of current evaluation practices for the comparative framework presented later in the paper.

A. Performance Measurement and Quality Improvement

Performance Measurement: Performance measurement is the process of collecting, analyzing, and reporting information regarding the performance of an individual, group, organization, system, or component. Started in the 1950s, this approach gained popularity with the Outcomes by Objectives and Performance Management movements. An early leader was William Edwards Deming, who developed an iterative four-step management method in the 1950's called "plan-do-check-act" (PDCA), also known as PDSA (plan-do-study-act), for performance management and continuous improvement of organizational processes and products. In 1987, the U.S. government introduced the concept of Total Quality Management and created the Malcolm Baldrige National Quality Award to promote quality achievements and publicize successful quality strategies. The original award criteria assessed financial, customer quality, internal process, and employee metrics.

In 1992, David Norton and Robert Kaplan published the Balanced Scorecard in an article for the Harvard Business Review. The Scorecard aligned financial customer, business process, and learning and growth (staff development) measures with a common corporate vision and strategy (Kaplan 2010). In 1993, the Government Performance and Results Act (GPRA) was enacted, which required federal agencies to develop and deploy a strategic plan, set performance targets, and measure their performance over time. More recently, the GPRA Modernization Act of 2010 (Public Law 111-352) was enacted with the goal of better integrating agency strategic plans, performance, and programs with more frequent reporting. These activities are meant to drive program improvement and are using performance data for more informed decision making.

Since the 1990s, more organizations have developed outcomes-based performance measurement and budgeting systems that use accountability-based evaluation methods to monitor program results and assess whether and to what extent program resources are managed well and attain intended results. Created to meet government- and funder reporting requirements, these evaluations track program process and outcome measures through management information systems.

Quality Improvement: Performance measures can be used for internal quality improvement processes within institutions and for external quality improvement processes across institutions. Continuous quality improvement (CQI) is an ongoing process to improve products and processes either through incremental improvement or sudden “breakthrough” improvement. In CQI, service delivery processes are evaluated in iterative cycles to improve their effectiveness and efficiency. The National Quality Measures Clearinghouse identifies three basic quality improvement steps: (1) identifying problems or opportunities for improvement, (2) selecting appropriate measures of these areas, and (3) obtaining a baseline assessment of current practices and then remeasuring to assess the effect of improvement efforts on performance (AHRQ 2013).

Baseline results can be used to (1) better understand a quality problem, (2) provide motivation for change, (3) establish a basis for comparison across institutional units or over time, and (4) enable prioritization of areas for quality improvement (AHRQ 2013). Although the genesis of quality improvement was in the healthcare industry, quality improvement methods can also be used to improve the implementation of evidence-based strategies for population health. For example, the Centers for Disease Control and Prevention (CDC) has specified a set of winnable battles: health priorities with large-scale impacts and known strategies for combating them (CDC 2013). The CDC is promoting the use of quality improvement methods through its Future of Public Health Awards to recognize positive initiatives in public health that use quality improvement to address winnable battles (Public Health Foundation 2012).

B. Rapid Cycle Evaluation

In 2010, the Patient Protection and Affordable Care Act (ACA) established the Center for Medicare and Medicaid Innovation (CMMI) to test innovative payment and service delivery models that aimed to improve the coordination, quality, and efficiency of care. The legislation provided \$10 billion in funding from 2011 to 2019 and enhanced authority to waive budget neutrality for testing new initiatives to allow quicker and more effective identification and spread of desirable innovations (Gold et al. 2011). CMMI created its Rapid Cycle Evaluation Group to evaluate the effectiveness of the new delivery and payment models. This group has created a new rapid cycle evaluation approach that will use summative and formative evaluation methods to rigorously evaluate the models’ quality of care and patient-level outcomes, while also delivering rapid cycle feedback to participating providers to help them continuously improve the models. The goal is to “evaluate each model regularly and frequently after implementation, allowing for the rapid identification of opportunities for course correction and improvement and timely action on that information” (Shrank 2013).

To maintain the rigor of its evaluations, the Rapid Cycle Evaluation Group is employing quasi-experimental designs that use repeated measures—time series analyses—to understand the relationship between implementation of new models and both immediate changes in outcomes and the rate of change of those outcomes. The group is also using other statistical methods, such as propensity score approaches and instrumental variables, and the use of comparison groups, where appropriate, to help clarify the models’ causal mechanisms. This approach assesses both the results and context of those results to gain a better understanding of how favorable outcomes are obtained. Evaluators collect qualitative information about (1) providers’ practices, organizational characteristics, (2) the culture of the health care systems in which they operate, (3) how providers implement the intervention, and (4) the factors that hinder and support the change.

This will allow evaluators to assess which features of the interventions are associated with successful outcomes (Shrank 2013).

Evaluators will submit these data to a CMMI Learning and Diffusion Team that has been organized to provide quarterly feedback to participating providers on dozens of performance metrics, including process, outcome, and cost measures. The team will also organize learning collaboratives among participating providers to “spread effective approaches and disseminate best practices, . . . ensuring that best practices are harvested and disseminated rapidly.” CMMI evaluation and dissemination activities are separated to “preserve the objectivity of the evaluation team” (Shrank 2013).

C. Complex Systems Evaluations

Innovators within and outside of government are also seeking to conduct systems change evaluations of large-scale, multisector, multilevel, and community-based initiatives, including those aimed at transforming juvenile justice systems, building healthy communities, and developing and implementing health reforms. Mathematica is currently evaluating these initiatives using a systems-based, mixed methods evaluation design that incorporates ongoing feedback. This three-part process matches the evaluation’s design with (1) the evaluation’s goals and intended uses, (2) the complexity of the intervention, and (3) the complex dynamics of the intervention’s context or environment (Hargreaves 2010). Other research and evaluation approaches have also been developed to provide rapid feedback to complex, large-scale initiatives: action research, developmental evaluation, and collective impact projects.

Action Research: Attributed to work done in the 1940s by Kurt Lewin, action research is an iterative research process involving researchers and community stakeholders. The approach creates a cycle of inquiry through three elements: “an ongoing analysis of contextual conditions, discrete actions taken to improve those conditions, and an assessment of the efficacy of those actions, followed by a reanalysis of the current conditions” (Foster-Fishman and Watson 2010). Action research has been used to evaluate change implemented at the organizational, inter-organizational, and community levels, including the expansion of health care services, the promotion of community development, the revision of education curricula, the creation of healthier communities, and the improvement of local government services (Springer 2007).

In the reflective or analytic phase of the inquiry cycle, the researcher facilitates a process in which the project stakeholders review the consequences of their actions and reflect on the effectiveness of their actions in solving the identified problem. This requires the creation of a group environment called *situated learning* that promotes dialogue and the development of a shared understanding of what new actions to take (Rosaen et al. 2001). Although the process theoretically ends when the original problem is solved, some argue that because the environment is constantly changing, this “cycle of inquiry, action, and reflection” can be used on a continuous basis (Rappaport 1981).

Recent adaptations of action research include participatory action research and systemic action research, which supports large-scale systemic change. Action research has more transformative, system-wide impact when it moves beyond first-order change. First-order action research projects make incremental improvements in existing programs and practices by asking how current practices can be done better, rather than asking why the system operates as it does.

In contrast, second-order, or transformative, action research projects ask why the current system operates as it does and seek to uncover the system's underlying patterns and the root causes of existing problems, which can result in a significant reframing of the issue and targeting of required changes to the system (Bartunek and Moch 1987).

Action research theories contributed to the work of Chris Argyris, and Don Schön, particularly on the concepts of single-loop and double-loop learning. Single-loop learning refers to an error and correction process that focuses on making a particular strategy more effective, without changing the underlying goals. In contrast, double-loop learning involves questioning the basic assumptions behind the goals and strategies, which leads to changes in the organization's underlying norms, policies, and objectives (Argyris 1982). Quality improvement is a form of single-loop learning. Frameworks for transformative systems change in action research use a more complex systems change approach that questions the system dynamics that led to the problem and intervenes in ways that modify underlying system relationships and functioning (Foster-Fishman and Watson 2010; Patton 2011).

Developmental Evaluation: Created by Michael Q. Patton, Developmental Evaluation (DE) applies complexity concepts to evaluation to support innovation development. Information collected through the evaluation is used to provide quick and credible feedback for adaptive and responsive development of an innovation. The evaluator works with the social innovator to cocreate the innovation through an engagement process that involves conceptualizing the social innovation throughout its development, generating inquiry questions, setting priorities for what to observe and track, collecting data, and interpreting the findings together to draw conclusions about next steps, including how to adapt the innovation in response to changing conditions, new learnings, and emerging patterns (Patton 2008, 2011). This evaluation approach involves double-loop learning, in which preliminary innovation theories and assumptions are reality tested and revised.

Developmental evaluation is designed to address specific dynamic contexts, including (1) the ongoing development and adaptation of a new program, strategy, policy, or initiative to new conditions in complex, dynamic environments; (2) the adaptation of effective principles to new contexts; (3) the development of a rapid response to a sudden major change or crisis, such as a natural disaster or economic meltdown; (4) the early development of an innovation into a more fully realized model; and (5) evaluations of major cross-scale, multilevel, multisector systems change (Patton 2011). The approach can use any kind of qualitative or quantitative data or design, including rapid evaluations. Methods can include, for example, surveys, focus groups, community indicators, organizational network analyses, consumer feedback, observations, and key informant interviews with influential community leaders or policymakers. The frequency of feedback is based on the nature and timing of the innovation. During slow periods, there is less data collection and feedback; when the initiative, such as a policy advocacy campaign, comes to a critical decision point, there is more frequent data collection and feedback (Patton 2011).

Collective Impact: Developed by John Kania and Mark Kramer of FSG, Collective Impact is a social change process for “highly-structured cross sector coalitions.” In Collective Impact projects, there is “heightened vigilance of multiple organizations looking for resources and innovations through the same lens, learning is continuous, and adoption happens simultaneously among many different organizations.” The approach is based on the implementation of five principles, or “conditions”: (1) the development of a common agenda involving a shared

understanding of the problem and a joint approach for solving it, (2) the consistent collection and measurement of results across all participants for mutual accountability, (3) the coordination of coalition participants' actions through a mutually reinforcing plan of action, (4) open and consistent communication among coalition participants, and (5) backbone support provided by an organization to coordinate the actions of participating organizations and agencies (Kania and Kramer 2013).

Not a rapid evaluation method per se, a Collective Impact project incorporates several action research, DE and systems change evaluation approaches. Collective Impact projects borrow methods from these other evaluation approaches, including the concepts of rapid learning, continuous feedback loops, the collective identification and adoption of new resources and solutions, the identification of patterns as they emerge, and an openness to unanticipated changes that would have fallen outside a predetermined logic model. To provide continuous feedback, the Collective Impact approach does not use episodic evaluation intended to assess the impact of discrete initiatives but promotes the use of complex systems-based evaluation approaches that are "particularly well-suited to dealing with complexity and emergence," such as DE (Kania and Kramer 2013). They also advocate creating ongoing feedback loops through weekly or biweekly reports, rather than the usual annual or semiannual evaluation timelines.

III. RAPID EVALUATION COMPARATIVE FRAMEWORK

It can be a challenge to identify which rapid evaluation methods are best suited for particular HHS studies. To allow for easier comparison of different rapid evaluation approaches, this framework identifies ten points of comparison: the intervention's (1) situational dynamics; (2) complexity; (3) governance structure; (4) scale of outcomes, (5) theory of change; (6) execution strategy; (7) sequence, scale, and timing of expected results; (8) evaluation purpose; (9) reporting and use of evaluation findings; and (10) evaluation methods. This section presents the details of this comparative framework for projects at three levels of complexity: quality improvement methods for simple process improvement projects, rapid cycle evaluations for complicated organizational change programs, and systems-based rapid feedback methods for large-scale systems change or population health initiatives.

Dynamic Complexity: All situations or systems share certain basic attributes or conditions, called boundaries, relationships, and perspectives. Together, these conditions generate patterns of system-wide behavior that are called situational or system dynamics.

- *Simple* dynamics are characterized by fixed, static, and mechanistic patterns of behavior, as well as linear, direct cause-and-effect relationships between system parts.
- In more *complicated* systems, leaders plan and coordinate the activities of multiple teams or parts. Because of circular, interlocking, and sometimes time-delayed relationships among units or organizations in complicated organizations, unexpected results can occur through indirect feedback processes.
- *Complex* adaptive dynamics are characterized by massively entangled webs of relationships, from which unpredicted outcomes emerge through the interactions of many parts or actors within and across levels. Complex systems are adaptive; actors

learn and coevolve as they interact with one another and respond to changes in their environment.

All dynamics are present to some degree within the same context or situation, although the balance of dynamics may shift over time (Patton 2010a). For example, hospitals have both complicated and complex dynamics; the dynamic that predominates depends on the context and organizational level.

Quality improvement, rapid cycle evaluation, and complex systems feedback approaches were designed for different interventions and system dynamics. Quality improvement approaches focus primarily on enhancing the efficiency of service delivery of specific practices within organizations. Rapid cycle evaluation methods were designed to evaluate organization-scale programs. Ongoing evaluation processes with rapid feedback mechanisms are appropriate for complex systems change initiatives. These are not mutually exclusive activities, however. They can be nested at appropriate layers, within initiatives. Each layer provides a context for the nested intervention.

Situational Fit: When is it appropriate to use each type of rapid evaluation method? The right design is one that best fits the evaluation's purpose(s) and captures the complexities of the intervention and its environment (Funnell and Rogers 2011). When system dynamics are not considered in an evaluation's design, the evaluation will inevitably miss crucial aspects of the intervention and its environment that are affecting the intervention's implementation, operation, and results. Factors to consider in selecting rapid evaluation methods include the dynamics of the intervention's context; the structure, program logic, and intended outcomes of the intervention itself; and the intended purpose and use of intervention's evaluation (Hargreaves 2010). To aid the matching process, these evaluation factors and methods are organized by level of complexity in a comparative framework (Table 1). More details of the evaluation factors are presented next.

Table 1. Rapid Evaluation Dynamics Comparative Framework

Evaluation Element	Evaluation Factor	Process Change	Organizational Change	Systems Change
Context	1. Situational dynamics	Simple	Complicated	Complex
	2. Type of intervention	Simple projects	Complicated programs	Complex initiatives
Intervention	3. Governance structure	Single organization	Federal funder of multiple grants	Alliance of multiple funders and stakeholders
	4. Scale of outcomes	Single, discrete process changes	Short list of individual-level outcomes	Large-scale population or system-wide change
	5. Timeline of expected results	Immediate change expected within weeks	Incremental change expected in months	Transformative change expected in months or years
	6. Theory of change	Implementing an evidence-based practice	Testing a specific program model	Applying change principles to strategic leverage points
	7. Execution strategy	Fidelity to a set of documented procedures	Fidelity to work plans outlining program goals, objectives, and strategies	Change strategies are developed and revised as the initiative evolves
Evaluation	8. Purpose	Implementation, efficacy, and outcome questions	Implementation, efficacy, and outcome questions	Implementation and efficacy questions
	9. Reporting and use of findings	Unit operations managers and staff receive and use evaluation results	Program management separates internal and external reporting and learning functions	Strategic leadership incorporates findings into adaptive management cycle
	10. Rapid evaluation methods	Quality improvement—plan-do-study-act cycle	Rapid cycle evaluation—formative and summative evaluation	Developmental evaluation, systems change evaluation, action research methods

A. The Context

Factor 1: Situational dynamics. How complex are the dynamics of the context or environment in which the intervention is operating?

Within the situation of interest, how complex are the set of relationships, connections, or exchanges among the players? In relatively simple situations, relationships are fixed, tightly

coupled (dependent), and often hierarchical. Complicated system relationships can involve the coordination of separate entities, such as the use of multiple engineering teams in designing and building a large-scale construction project (Gawande 2010). In contrast, complex system dynamics are characterized by less centralized control.

How different are the stakeholders' perspectives in the situation? In simple contexts, there is a high degree of certainty and agreement over system goals and how to achieve them; there is consensus on both desired ends and means (Stacey 1993). In complicated situations, there may be agreement on the overall purpose or goal of a program or intervention, but less certainty and consensus on how to achieve it. For example, neighbors may come together to advocate for improved traffic safety but prefer a range of solutions, from sidewalks to stop signs. In complex situations, such as education reform, there may be a great diversity of perspectives regarding both reform goals and strategies among the families, teachers, school administrators, education board members, state legislators, federal officials, and other stakeholders involved in the issue.

B. The Intervention

Factor 2: Intervention complexity. Is the intervention a simple, direct process change, a test of a program model, or a larger initiative addressing multisector, multilevel population or systems change?

Factor 3: Governance structure. Who is funding and overseeing the initiative—a single organization, a federal funder of a cohort of grantees, or a consortium of funders?

How complicated is the program, including its structure? An intervention's governance may include its funding, management, organizational structure, and implementation. A simple intervention is typically implemented by a single organizational unit. More-complicated efforts involve teams of experts or program units within an organization. Complex, networked interventions often involve the collaboration of multiple entities or organizations working across sectors and levels. Examples include comprehensive public health initiatives addressing childhood asthma and obesity; community capacity building initiatives to prevent child abuse and neglect; integrated urban development initiatives alleviating poverty; and networks of federal, state, and local agencies working together on homeland security (Kamarck 2007; Goldsmith and Kettl 2009).

Factor 4: Scale of outcomes. What is the expected change—a process improvement, changes in a small, specified set of individual-level outcomes, or broader system-wide change?

Factor 5: Timeline of expected results. Will early results be seen immediately or in weeks, months, or years?

What is the scope of the program's desired results? Simple, linear interventions are designed to produce specific, narrowly focused, and measurable outcomes, such as increasing student reading skills through a literacy program, or increasing the accuracy of the results of lab tests through changes in testing procedures. Interventions with more-complicated dynamics may target multiple, potentially conflicting outcomes, such as improving personal safety while maintaining the quality of life and level of independence for seniors in community-based, long-term care programs (Brown et al. 2008). In complex system interventions, stakeholders may

share a common vision, such as reduced poverty at a regional level, improved quality of life for people with developmental disabilities, or rapid advances in biomedical science, but they may not be able to predict in advance which specific outcomes will emerge from the complex interactions of the initiative's entities or organizations.

Factor 6: Theory of change. Is the program developing, adapting, or implementing a promising or best practice, testing a program model, or applying general principles to a complex systems change process?

How complex are the program's dynamics? In simple, straightforward interventions, linear logic models can be used to trace a single stream of program inputs, activities, and outputs that lead to a limited set of outcomes. In more-complicated interventions, multiple coordinated pathways of activities may lead to a broader set of complementary outcomes. Systems change interventions may target specific drivers of systemic change at strategic leverage points. Potential drivers might include the development of a shared vision, collaborative partnerships, pooled resources, and shifts in community norms. The activities used to facilitate those changes might include convening meetings of potential partners or changing policies to allow the pooling of resources.

Factor 7: Execution strategy. How prescribed is the intervention's implementation strategy? Is the implementation work plan a clearly specified set of procedures, a program manual, set of program guidelines, or contract requirements outlining the specific functions, timing, and sequence of program components, or is the execution strategy developed collectively and adapted over time with input from the initiative's partners and stakeholders?

C. The Evaluation

Factor 8: Purpose. What are the goals of the evaluation? Are the evaluation questions about implementing and testing the efficacy of a particular best practice or program model in a specific context, or making a judgment of the program's value? Or, are the questions addressing how best to move forward in a complex initiative?

Factor 9: Reporting and use of findings. When, how, and to whom are results reported? Is reporting linked to, or kept separate from, sessions with decision makers and stakeholders to understand and interpret the findings and take action in response?

Factor 10: Rapid evaluation methods. Which evaluation methods are the best match for the circumstances?

Because evaluation designs depend on the kinds of evaluation questions asked as well as on the system conditions and dynamics, there is no one best design option. The right design is one that addresses the evaluation's purpose(s) and captures the complexities of the intervention and its context or environment. Rapid evaluation methods can be used in developmental, formative, or summative evaluations, addressing questions regarding the implementation of a process improvement, program, or larger initiative, how it can be improved, and its cost-effectiveness. Quality improvement, rapid cycle, and systems change evaluation approaches can use similar qualitative and quantitative research methods, including feedback surveys, focus groups, key informant interviews, and tracking of performance indicators.

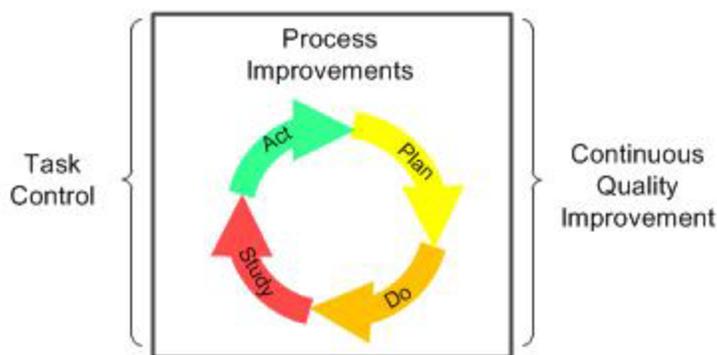
One element that sets apart the different rapid evaluation designs is their feedback mechanism, including how their findings are reported, to whom, and for what purposes. Although the results of quality improvement projects can be disseminated broadly, the target audience for the findings is the internal program unit whose processes are being changed. External funders are a key audience for the findings of rapid cycle evaluations, although the findings are also reported back to the organizations implementing the grant or program model. In rapid cycle evaluations, the evaluation’s cross-site findings are reviewed and interpreted at the funder level, not at the grantee level, to maintain the evaluation’s objectivity. In contrast, in complex initiatives the lines between internal and external evaluation audiences are blurred. In complex initiatives, collaborative learning processes might be used to convene initiative leaders and stakeholders to learn about, understand, and interpret the results, and to make collective decisions about how to improve and adapt the initiative based on the evaluation findings.

In the next three sections (Section IV. V, and VI), rapid evaluation examples for three types of change initiatives, process change, organizational change, and systems change, illustrate the match between evaluation design, contextual complexity, and content of the intervention. For each example, the ten evaluation factors highlighted in Table 1 are described in more detail in the section-specific tables.

IV. PROCESS CHANGE—QUALITY IMPROVEMENT

This section provides an example of a process change evaluation used in the implementation of the Safe Surgery Checklist. Table 2 features the unique factors of process change evaluations. In Table 2, the ten evaluation factors described in the previous section are listed (in the left-hand column), the factors are applied to process change projects (in the middle column), and the factors are illustrated through one surgery checklist example (in the right-hand column). The rapid evaluation practice of process improvement is illustrated in Figure 1.

Figure 1. Continuous Quality Improvement



Quality improvement methods are appropriate rapid evaluation techniques for simple, discrete process improvements in the daily functions of organizational units). Dr. Atul Gawande demonstrated the value of this technique in the World Health Organization’s Safe Surgery Checklist Project, when his team used quality improvement methods to test early prototypes of a surgical checklist (Haynes et al. 2009, Gawande 2010). The checklist was designed to reduce preventable surgical errors.

Through numerous PDSA cycles of prototype testing, the team improved the checklist, reducing the number of items, limiting the checklist to one page, and streamlining its application to two minutes. At that point, the team implemented a pilot study, in which surgical teams in eight hospitals adapted and implemented the checklist in their operating rooms. Time series analyses from the pilot study showed that surgical complications and deaths were reduced significantly (36 and 47 percent respectively) after the checklists were introduced (Gawande 2010).

Although many hospitals have implemented the checklist, it is not yet being used to its full capacity. By the end of 2009, approximately 10 percent of American hospitals had adopted the checklist or had started taking steps to implement it. Globally, more than 2,000 hospitals had started using the checklist. The uptake of the checklist has been slower than expected, however. Gawande recognized that on its own, a single process change such as the checklist cannot institute a complex, system-wide culture shift among hospitals and physicians to test and improve their surgical practices:

Just ticking boxes is not the ultimate goal here. Embracing a culture of teamwork and discipline is. And if we recognize the opportunity, the two-minute World Health Organization checklist is just a start. It is a single, broad-brush device, intended to catch a few problems common to all operations, and we surgeons could build on it to do even more (Gawande 2010).

Through iterative testing of adaptations of the Safe Surgery Checklist, other hospitals have improved its practice. Some researchers have found that the introduction of the hospital checklist initially lowered the risk of mistakes, but then the error rate gradually returned back to near its former level. To address this problem, one surgeon, Marc Parnes, tested an adaptation of the checklist by having a “personal check-in conversation” with the patient while rolling the patient into the procedure room. The conversation with the patient and entire operating team allowed each person to see the situation through the eyes of the others, including the patient. This adaptation reduced the hospital’s surgical error rate more sustainably than did the original checklist (Scharmer and Kaufer 2013).

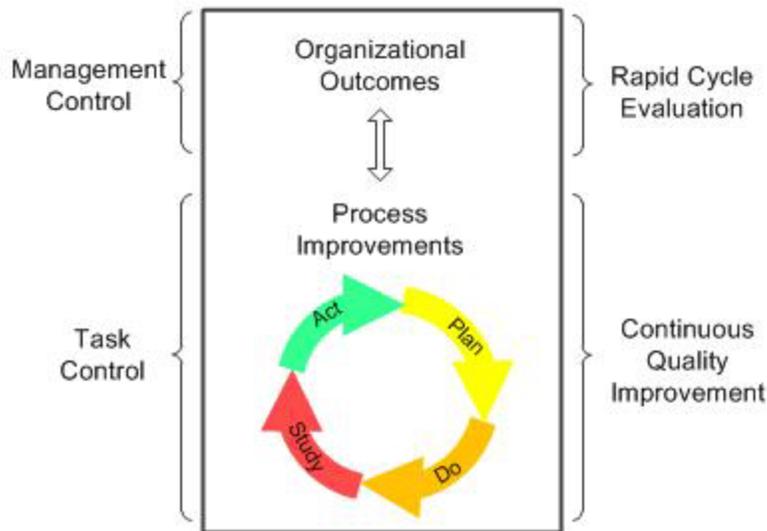
Table 2. Rapid Evaluation: Simple Process Change

Evaluation Factor	Process Change	World Health Organization's Safe Surgery Checklist
1. Situational dynamics	Simple	The team's dynamics are simple; the surgeon leads the surgical team. In surgery, the patient's unstable health adds some complexity to the situation.
2. Intervention complexity	Simple projects	Simple project: the surgical team completes a two-minute, 19-step checklist designed to prepare teams better for surgery and to respond better to unexpected problems that occur during surgery.
3. Governance structure	Organizational unit	Hospital operating room teams implement the checklist.
4. Scale of outcomes	Single, discrete process changes	The checklist changes the quality of the procedures used to prepare the patient and the surgical equipment before surgery and to prepare the team for potential problems during the operation.
5. Timeline of expected results	Immediate change expected within weeks	The checklist was introduced to operating rooms over a period of one week to one month. Data collection started during the first week of checklist use.
6. Theory of change	Implementing an evidence-based practice	The same procedure was used in all situations, with minor adaptations in language, terminology, and the order of the checklist items for different hospitals.
7. Execution strategy	Fidelity to a set of documented procedures	Hospital surgical teams received the checklist with a set of how-to PowerPoint slides and YouTube videos.
8. Purpose	Implementation, efficacy, and outcome questions	The eight-hospital pilot study tested the checklist's efficacy; major complications were reduced by 36 percent and deaths were reduced by 46 percent after the introduction of the checklist.
9. Reporting and use of findings	Unit operations managers and staff receive and use evaluation results	To create the prototype of the checklist, Gawande's surgical team tested multiple versions of the original checklist using operation process and outcome data, and team feedback, to track improvement of the checklist.
10. Rapid evaluation methods	Quality improvement—PDSA cycle	The team developing the checklist used the results of each PDSA cycle to improve the checklist, so that it was ready to be tested in operating rooms in other hospitals.

V. ORGANIZATIONAL CHANGE—RAPID CYCLE EVALUATION

This section provides an example of an organizational change evaluation used in the implementation of the Partnerships for Patients Program (PfP). Table 3 features the unique factors of organizational change evaluations. In Table 3, the ten evaluation factors are listed (in the left-hand column), the factors are applied to organizational change projects (in the middle column), and the factors are illustrated through the PfP example (in the right-hand column). The practice of rapid cycle evaluation is illustrated in Figure 2.

Figure 2. Rapid Cycle Evaluation



Rapid cycle evaluations are appropriate for testing models of organizational change. The Center for Medicare and Medicaid Innovation (CMMI) developed the rapid cycle evaluation approach to test innovative health care payment and service delivery models that preserve or improve the quality of care while reducing costs (HHS 2011, Shrank 2013). The Centers for Medicare & Medicaid Services (CMS) is using this approach to evaluate several national initiatives, including PfP. PfP aims to prevent hospital-acquired infections and hospital complications, increasing patient safety and cutting related hospital readmissions. More than 3,700 hospitals are participating in the initiative (CMS 2013a). To reduce preventable inpatient harms by 40 percent and readmissions by 20 percent by the end of 2013, the PfP hospitals need to invest in and redesign their organization infrastructures to support improved care. This requires “substantial learning and adaptation” on the part of health care providers as “there are no simple turnkey solutions” (Shrank 2013).

To help hospitals identify and implement effective solutions, CMS awarded \$218 million in 2011 to 26 state, regional, national, or hospital system organizations to become Hospital Engagement Networks (HENs).¹ The HENs support the initiative by identifying hospitals’

¹ Since then, one more HEN has been awarded, increasing the number of HENs to 27.

current solutions to reducing hospital-acquired conditions and disseminating them to other hospitals and health care providers. The HENs are using a range of strategies to help hospitals, including providing financial incentives, developing learning collaboratives, conducting intensive training programs, providing technical assistance, implementing data systems to monitor hospitals' progress, and identifying high-performing hospitals (CMS 2013b). The HENs are required to develop, collect, and report PfP process and outcome measures that monitor the early progress of the initiative. The timing, content, and quality of the HEN data have varied considerably (Felt-Lisk 2013).

Rapid cycle evaluation methods are being used in PfP's formative implementation and summative impact evaluations that are being conducted by the team of Health Services Advisory Group, Inc. (HSAG) and Mathematica. The evaluations' goals are to provide real-time information monitoring the hospitals' progress, and ongoing feedback to support improvement of PfP activities and outcomes. Specifically, the formative evaluation is (1) documenting the organizational context of the PfP hospitals, including the hospitals' infrastructure (staffing and operational systems) and level of commitment to the initiative's 11 clinical areas of focus; (2) monitoring hospitals' site-specific process measures that are reported monthly by their HENs; and (3) documenting hospitals' activities and challenges in monthly HEN reports to CMS. Whether PfP's goals are met will be determined by time series analyses of pre-post patient data derived from medical chart reviews conducted by the Agency for Healthcare Research and Quality, using a sample representative of the entire nation. In addition, an impact evaluation conducted by the HSAG/Mathematica team will determine the extent to which changes over time can be attributed to PfP and will identify if there are certain types of interventions associated with greater harm reduction (Felt-Lisk 2013).

Key internal audiences for the evaluation findings are CMS staff, support contractors, and the leadership and staff of the HENs, all of whom have a voracious appetite for information (Felt-Lisk 2013). Monthly Formative Feedback Reports include a one-page visual summary of the initiative's progress, key news, appendices for evaluation methods and supplemental tables organized by clinical area (200+ pages) and by HEN (200+ pages). CMS has also made special requests for graphs of individual site-level (hospital) progress and for "success stories" of downward trends in preventable harms, accompanied by a description of associated interventions.

There are additional costs associated with being able to "provide formative information to feed program needs at any given point in concert with the flow of the program" beyond what is done in traditional evaluation. New reporting methods are also required for the impact evaluation, with frequently updated deliverables that are timed as soon as data are available, using PowerPoint presentations and simple memos with tables, rather than formal reports. As a result, the program's formative evaluator noted that while rapid cycle evaluation's methods might be somewhat more costly than traditional evaluation, the approach appeared to be more useful to the evaluation's primary audiences than other approaches (Felt-Lisk, personal communication, Sept. 2013).

Table 3. Rapid Evaluation: Complicated Organizational Change

Evaluation Factor	Organizational Change	CMS Partnership For Patients Campaign
1. Situational dynamics	Complicated	Complicated campaign, with some complexity of learning across HENs in an “all teach and all learn” environment.
2. Intervention complexity	Complicated programs	National initiative to reduce preventable patient harms in 3700 hospitals across 27 HENs.
3. Governance structure	Federal funder of multiple grants	CMS is the single federal funder of the PfP.
4. Scale of outcomes	Short list of individual-level outcomes	PfP has two overarching outcomes (reduced hospital-acquired patient harms, reported in 11 clinical areas, and hospital readmissions).
5. Timeline of expected results	Incremental change expected in months	Results are expected in months; the focus is on speeding up the pace of change to achieve the goals in three years.
6. Theory of change	Testing a specific program model	HENs facilitate sharing of best practices among their aligned hospitals and offer hospitals training, technical assistance, learning collaboratives, and reporting systems to help them achieve the PfP goals.
7. Execution strategy	Fidelity to work plans outlining program goals, objectives, and strategies	Detailed hospital-specific work plans and measures of success are developed and implemented by the hospitals.
8. Purpose	Implementation and efficacy questions	In what contexts are the PfP hospitals working to achieve the PfP goals? What progress are the hospitals making? What are early results? How do PfP implementation and early outcomes vary by hospital, HEN, and condition?
9. Reporting and use of findings	Program management separates reporting and learning functions	Monthly feedback report to CMS and HENs, and ad hoc reports in response to special requests. No external evaluation linkage to CMS’s internal Learning Team to maintain objectivity.
10. Rapid evaluation methods	Rapid cycle formative and summative evaluation methods	Rapid cycle formative and summative evaluation methods; time-series analyses of changes in hospital-specific processes and outcomes, with documentation of hospitals’ contexts, culture, and PfP activities.

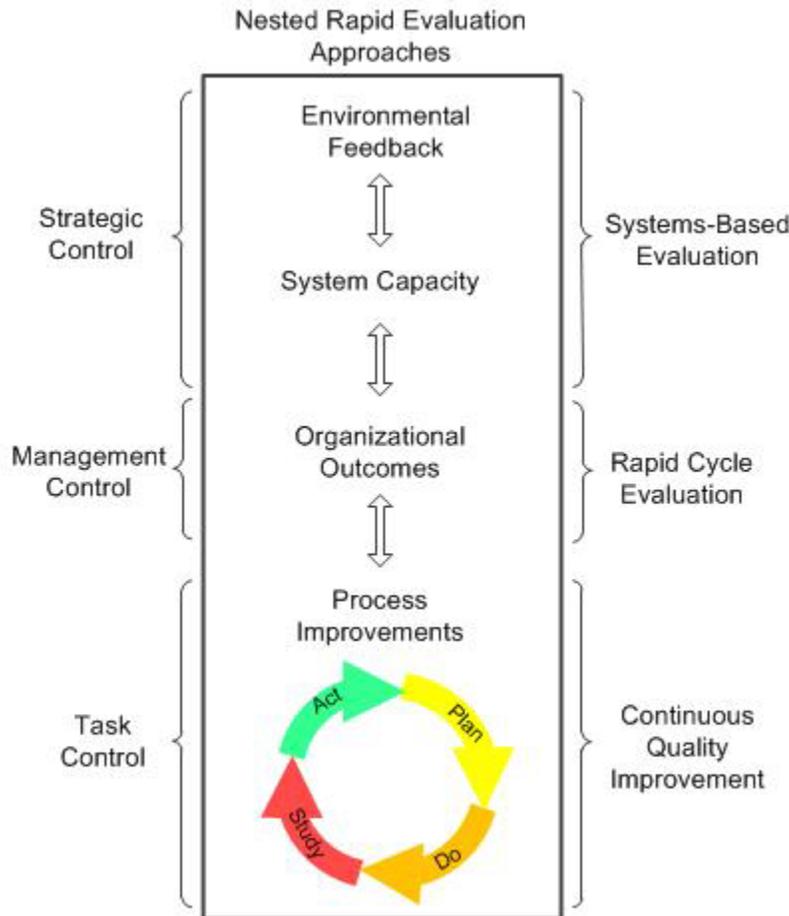
CMS = Centers for Medicare & Medicaid Services; HEN = Hospital Engagement Network; PfP = Partnership for Patients

VI. COMPLEX SYSTEMS CHANGE—SYSTEMS CHANGE EVALUATION

This section provides an example of a system change evaluation used in the implementation of the Medicaid and CHIP Learning Collaboratives (MAC LC) project. Table 4 features the unique factors of system change evaluations. In Table 4, the ten evaluation factors are listed (in the left-hand column), the factors are applied to systemic change projects (in the middle column),

and the factors are illustrated through the MAC LC example (in the right-hand column). The rapid evaluation practice of systems change evaluation is illustrated in Figure 3.

Figure 3. Nested, Systems-based Evaluation



There are evaluation methods designed for initiatives that engage many different entities in collaborative efforts to shift larger systems. The system-shifting strategies in those initiatives are not predetermined best practices, or fully specified program models (although they may incorporate both). Their strategies are developed over time as key actors work across organizations, sectors, and levels to achieve common goals. Development-oriented, systems-based evaluation methods with rapid feedback mechanisms were created to address these systems change initiatives. These evaluation methods differ from quality improvement and rapid cycle evaluation methods in four ways. First, they use nested logic models and theories of change that show how the interactions of unit (micro)-, organization (meso)-, and policy or community (macro)-level activities impact individual-level outcomes. Second, they recognize, document, and incorporate the changing dynamics of the initiative’s environment into the evaluation. Third, they include systems concepts in the evaluation’s conceptual framework to measure changes in collective capacity, networked relationships, and shared perspectives of the entities and organizations involved in the initiative. Fourth, they embed the evaluation function into the initiative as an integral set of feedback loops to support decision making at tactical unit, program management, and strategic initiative levels.

This development-oriented, systems-based approach is currently being used by CMS in the MAC LC project, established in 2011 to achieve high-performing state health coverage programs, a goal that requires “a robust working relationship among federal and state partners.” The original two-year MAC LC project brought federal and state Medicaid agencies together “to address common challenges and pursue innovations in Medicaid program design and operations as well as broader state health coverage efforts” (CMS 2013c). The project created six collaborative work groups, each consisting of 6 to 10 states plus federal partners and national experts, which addressed a range of topics “critical for establishing a solid health insurance infrastructure,” including policies related to the implementation of the ACA. The original six learning collaboratives (LCs) were the (1) Exchange Innovators in Information Technology LC, (2) Expanding Coverage LC, (3) Federally Facilitated Marketplace Eligibility and Enrollment LC, (4) Data Analytics LC, (5) Promoting Efficient and Effective IT Practices LC, and (6) Value-Based Purchasing LC.²

The MAC Collaboratives activities are coordinated by Mathematica, the Center for Health Care Strategies (CHCS), and Manatt Health Solutions, with additional assistance from external experts and in close association with CMS. Over a period of two years, LC meetings, called learning sessions, were conducted, mostly by webinar or conference call, on a monthly or biweekly basis, and were moderated by Mathematica, CHCS, and Manatt facilitation teams. In the learning sessions, state representatives, technical experts, and CMS staff discussed policy issues, reviewed draft rules and other federal guidance, and created technical assistance tools, background materials, and other state resources (CMS 2013d).

The MAC LC project included an internal assessment function operating at three levels (learning session, LC, and project), using a systems-based, multilevel conceptual framework. First, the assessment team observed and rated the quality of the content, logistics, and facilitation of individual learning sessions, for quality improvement purposes. Second, the assessment team tracked LC session attendance rates, conducted participant feedback surveys, and reviewed LC documents to evaluate the performance of each LC, for formative purposes. Third, the assessment team also interviewed the CMS staff and LC facilitation teams to obtain project-level information about the overall functioning and effectiveness of the project. Assessment feedback is provided on different cycles, through monthly debriefings with the project director on the LC sessions, group feedback to the LC facilitation teams in quarterly project management meetings, and annual reports to CMS on the performance of the project as a whole. This feedback is intended primarily for an internal audience.

In 2013, the project was renewed for two more years; some aspects of the project were modified, informed by internal and external feedback. Over the next two years, the assessment team will continue to provide an internal monitoring and rapid feedback function on the project. There are costs associated with this function; the internal assessment is a separate project task.

² In 2013, the Promoting Efficient and Effective IT Practices LC ended and was replaced by the Basic Health Plan LC.

Table 4. Rapid Evaluation: Complex Systems Change

Evaluation Factor	Systems Change	CMS Medicaid and CHIP Learning Collaborative
1. Situational dynamics	Complex	There are complex dynamics between CMS and states and among CMS, states, and facilitators; the project is also operating within the volatile political dynamics of federal health care reform.
2. Intervention complexity	Complex initiatives	Three organizations contracted by CMS to operate six learning collaboratives addressing ACA and non-ACA topics.
3. Governance structure	Alliance of multiple funders and stakeholders	CMS funds LC facilitation teams and provides in-kind CMS expertise; 50 states voluntarily participate in one or more LCs by invitation.
4. Scale of outcomes	Large-scale population or system-wide change	Balance of short-term outcomes (LC-created policies, tools, and practices) and long-term outcomes, such as successful implementation of the ACA.
5. Timeline of expected results	Transformative change expected in months or years	The coordination of state and federal learning around Medicaid policy is ongoing. Several initiatives, such as the implementation of the ACA, will require several more years of implementation.
6. Theory of change	Applying change principles to strategic leverage points	The LCs provide a new forum for federal-state dialogue, creating new communication channels that increase federal-state collaboration on critical issues.
7. Execution strategy	Change strategies developed and revised as initiative evolves	The content, format, frequency, and facilitation of the learning sessions were modified over time in response to participant feedback and changes in the federal policy landscape.
8. Purpose	Implementation and efficacy questions	Developmental and formative questions: What are the LCs doing to develop and implement learning sessions that meet the needs of CMS and state representatives? How can the structure and process of the LCs be improved?
9. Reporting and use of findings	Strategic leadership incorporates findings into adaptive management cycle	Monthly project direct debriefings, quarterly project management presentations, annual performance reports to CMS—no external publication of results.
10. Rapid evaluation methods	Developmental evaluation, systems change evaluation, action research methods	Direct observation of sessions, immediate post-session web surveys of state participants, monitoring of session attendance metrics, in-depth interviews with facilitation members and CMS staff, and ongoing review of project documentation.

CMS = Centers for Medicare & Medicaid Services; CHIP = Children’s Health Insurance Program; ACA = Patient Protection and Affordable Care Act; LC = learning collaborative

VII. CONCLUSIONS

This comparative framework is a heuristic tool rather than a prescriptive how-to manual for assigning rapid evaluation methods to different projects. There is not one best rapid evaluation method that works in all circumstances. The right rapid evaluation design addresses the goals of the evaluation and captures the complexities of the intervention and its environment. However, these methods are not mutually exclusive. They may be more effective when nested, just as the basic building blocks of genetic code are used in different combinations to develop biological functions that work together to form organisms that evolve over time (Holland 1995). Or, as illustrated in this paper, simple checklists are used to reduce preventable surgical errors within larger hospital safety campaigns that are funded through national payment reforms that reward system-wide shifts in health care costs and quality. Evaluating an intervention from process, organization and system perspectives allows us to implement change more effectively from multiple vantage points.

Several universal rapid evaluation principles are important, regardless of the level of complexity. First, rapid evaluation methods should maintain a balance between short-term results and long-term outcomes, so that there is “an alignment of task control, management control and strategic control” (Kaplan 2010). In other words, a short-sighted emphasis on immediate results (the optimization of subsystems), should not jeopardize the achievement of long-term goals (the optimization of the whole system). Second, rapid evaluation should not just be a measurement or diagnostic tool; it should also be part of an interactive and adaptive management process, in which internal operational results and external environmental feedback are used together in an iterative process to test and revise an initiative’s overall strategy. Third, the information collected, analyzed, and interpreted should be used “as a catalyst for continual change,” in which data and action plans are reconsidered and original assumptions are questioned through a reflective, double-loop learning process that supports rethinking of project goals (doing the right thing) as well as project strategies (doing things right) (Argyris 1982).

As federal initiatives become more complex and pressures increase to learn quickly from them, there will be many more opportunities to use these methods, alone or in combination, to improve the effectiveness of a wide range of initiatives. For example, this framework can be used within HHS for complex public health and human service programs as well as for broader systemic reforms. The most valuable rapid evaluation projects may include a combination of rapid evaluation approaches.

BIBLIOGRAPHY

- Argyris, C. *Reasoning, Learning, and Action: Individual and Organizational*. San Francisco: Jossey-Bass, 1982.
- AHRQ. (Agency for Healthcare Research and Quality). National Quality Measures Clearinghouse. “Uses of Quality Measures.” Available at [<http://www.qualitymeasures.ahrq.gov/tutorial/using.aspx>]. Accessed September 20, 2013.
- Bartunek, J. M., and M. K Moch. “First-Order, Second-Order, and Third-Order Change and Organizational Development Interventions: A Cognitive Approach.” *Journal of Applied Behavioral Science*, vol. 23, no. 4, 1987, pp. 483–500.
- Brown, Randall S., Carol V. Irvin, Debra J. Lipson, Samuel E. Simon, and Audra T. Wenzlow. “Research Design Report for the Evaluation of the Money Follows the Person (MFP) Grant Program.” Final report submitted to the Centers for Medicare & Medicaid Services. Princeton, NJ: Mathematica Policy Research, October 3, 2008.
- CDC. (Centers for Disease Control and Prevention.) “Winnable Battles.” Available at [<http://www.phf.org/programs/winnablebattles/Pages/default.aspx>]. Accessed September 20, 2013.
- CMS. (Centers for Medicare and Medicaid Services.) 2013a. “Partnerships for Patients Initiative.” Available at [<http://innovation.cms.gov/initiatives/partnership-for-patients/>]. Accessed September 18, 2013.
- . 2013b. “Hospital Engagement Networks.” Available at [<http://partnershipforpatients.cms.gov/about-the-partnership/hospital-engagement-networks/thehospitalengagementnetworks.html>]. Accessed September 19, 2013.
- . 2013c. “Medicaid and CHIP (MAC) Learning Collaboratives.” Available at [<http://www.medicaid.gov/State-Resource-Center/MAC-Learning-Collaboratives/Medicaid-and-CHIP-Learning-Collab.html>]. Accessed September 18, 2013.
- . 2013d. “MAC Collaboratives: State Tool Box.” [<http://www.medicaid.gov/State-Resource-Center/MAC-Learning-Collaboratives/Learning-Collaborative-State-Toolbox/MAC-Collaboratives-State-Toolbox.html>]. Accessed September 18, 2013.
- Eoyang, G. “Human Systems Dynamics: Complexity-Based Approach to a Complex Evaluation.” In *Systems Concepts in Evaluation: An Expert Anthology*, edited by Bob Williams and Iraj Imam. Point Reyes Station, CA: EdgePress. American Evaluation Association, 2007.
- Felt-Lisk, S. “Techniques to Support Rapid Cycle Evaluation: Insights from the Partnership for Patients Evaluation.” Presented at the AcademyHealth Annual Research Meeting, Baltimore, MD, June 25, 2013.

- Foster-Fishman P., and E. Watson. "Action Research as Systems Change." In *Handbook of Engaged Scholarship: The Contemporary Landscape. Volume Two: Community-Campus Partnerships*, edited by H. E. Fitzgerald, D. L. Zimmerman, C. Burack, and S. Seifer. East Lansing, MI: Michigan State University Press, 2010.
- Funnell, S., and P. Rogers. *Purposeful Program Theory: Using Theories of Change and Logic Models to Envisage, Facilitate, and Evaluate Change*. San Francisco: Jossey-Bass, 2011.
- Gawande, A. *The Checklist Manifesto: How to Get Things Right*. New York: Metropolitan Books, 2010.
- Gold, M., Helms, D., and S. Guterman. "Identifying, Monitoring, and Assessing Promising Innovations: Using Evaluation to Support Rapid-Cycle Change." Commonwealth Fund Publication 1512, Vol. 12. New York: The Commonwealth Fund, 2011.
- Goldsmith, S., and D. Kettl (eds.). *Unlocking the Power of Networks: Keys to High-Performance Government*. Washington, DC: Brookings Institution, 2009.
- Hargreaves, M. "Evaluating System Change: A Planning Guide." Cambridge, MA: Mathematica Policy Research, April 2010.
- Hargreaves, M., T. Honeycutt, C. Orfield, M. Vine, C. Cabili, M. Morzuch, S. Fisher, and R. Briefel. "The Healthy Weight Collaborative: Using Learning Collaboratives to Enhance Community-Based Initiatives Addressing Childhood Obesity." *Journal of Health Care for the Poor and Underserved*, vol. 24, no. 2, suppl., May 2013, pp. 103–115.
- Haynes, A., W. Berry, S. Lipsitz, A. Hadi, S. Breizat, E. Patchen Dellinger, T. Herbosa, S. Joseph, P. Kibatala, M. Lapitan, A. Merry, K. Moorthy, R. Reznick, B. Taylor, and A. Gawande. "A Surgical Safety Checklist to Reduce Morbidity and Mortality in a Global Population." *New England Journal of Medicine*, vol. 360, January 2009, pp. 491–499.
- HHS. (United States Department of Health and Human Services.) "Partnership for Patients Launched to Improve Care and Lower Costs for Americans." News release, June 2011. [<http://www.ahrq.gov/news/newsletters/research-activities/jun11/0611RA30.html>]. Accessed March 31, 2014.
- Holland, J. H. *Hidden Order: How Adaptation Builds Complexity*. Reading, MA: Helix Books, 1995.
- Kamarck, Elaine C. *The End of Government . . . As We Know It: Making Public Policy Work*. Boulder, CO: Lynne Rienner Publishers, 2007.
- Kania, J., and M. Kramer. "Embracing Emergence: How Collective Impact Addresses Complexity." Blog entry, January 21, 2013. Stanford Social Innovation Review. Available at [http://www.ssireview.org/blog/entry/embracing_emergence_how_collective_impact_addresses_complexity].

- Kaplan, R. S. "Conceptual Foundations of the Balanced Scorecard." Working Paper 10-074. Cambridge, MA: Harvard Business School, 2010. Available at [<http://www.hbs.edu/faculty/Publication%20Files/10-074.pdf>]. Accessed September 20, 2013.
- Patton, M. Q. *Utilization-Focused Evaluation*, 4th ed. Thousand Oaks, CA: Sage Publication, 2008.
- . *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: The Guilford Press, 2011.
- Perla, R. J., L. Provost, and G. Parry. "Seven Propositions of the Science of Improvement: Exploring Foundations." *Quality Management in Health Care*, vol. 22, no. 3, July-September 2013.
- Public Health Foundation. "Future of Public Health Awards." News release, May30, 2012. Available at [http://www.phf.org/news/Pages/Public_Health_Foundation_Names_Future_of_Public_Health_Award_Recipients.aspx]. Accessed September 18, 2013.
- Rappaport, J. "In Praise of Paradox: A Social Policy of Empowerment Over Prevention." *American Journal of Community Psychology*, vol. 9, no. 1, February 1981, p.1.
- Rosaen, C. L., P. Foster-Fishman, and F. Fear. "The Citizen Scholar: Joining Voices and Values in the Engagement Interface." *Metropolitan Universities*, vol. 12, no. 4, December 2001, pp. 10–29.
- Scharmer, O., and K. Kaufer. *Leading from the Emerging Future: From Ego-system to Ecosystem Economies*, San Francisco, CA: Berrett-Koehler Publishers, Inc., 2013.
- Shrank, W. "The Center for Medicare and Medicaid Innovation's Blueprint for Rapid-Cycle Evaluation of New Care and Payment Models." *Health Affairs*, vol. 32, no. 4, April 2013, pp. 807–812.
- Springer, E. *Action Research*, 3rd ed. Los Angeles, CA: Sage Publications, 2007.
- Stacey, R. *Strategic Management and Organizational Dynamics: The Challenge of Complexity*. Essex, England: Financial Times Prentice Hall, 1993.

MATHEMATICA
Policy Research

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and surveys

Princeton, NJ ■ Ann Arbor, MI ■ Cambridge, MA ■ Chicago, IL ■ Oakland, CA ■ Washington, DC

Mathematica® is a registered trademark of Mathematica Policy Research