



Improving the Science of Continuous Quality Improvement Program and Evaluation

An RWJF national program

SUMMARY

Improving the Science of Continuous Quality Improvement Program and Evaluation funded teams of researchers to address a core question within health care environments: “How will we know that a change is an improvement?” Research teams tackled an array of projects aimed at improving evaluation frameworks, quality improvement measures, and data collection and methodology. The Robert Wood Johnson Foundation (RWJF) authorized the program for up to \$1.5 million for 48 months, from August 2007 through August 2011.

CONTEXT

The health care field, in pursuit of better care processes and outcomes, has adopted techniques from industry, particularly continuous quality improvement or CQI. CQI is the process-based, data-driven approach to improving the quality of a product or service through iterative cycles of action and evaluation. However, the CQI field has had limited success in producing effective tools, strategies, or insights to improve quality in health care, despite strong interest in these techniques within the health care field.

RWJF’s Interest in This Area

RWJF has funded several national programs that employed CQI approaches to improve health care processes, staff engagement, and patient outcomes.

- *Improving Chronic Illness Care*, 1998 to 2010, aimed to improve the health of chronically ill patients by helping large numbers of organized health systems redesign how they deliver care through the Chronic Care Model. The program’s clinical improvement program engaged dozens of practices and health systems throughout the country in breakthrough collaboratives. For more information, read the [Program Results Report](#).

- *Pursuing Perfection*, 2001 to 2008, supported efforts by seven health care organizations to dramatically improve their care processes and patient outcomes, and to demonstrate to the broader provider community that ideal care is attainable. The program used CQI tools such as Plan-Do-Study-Act (PDSA) cycles¹ and improvement collaboratives to accomplish its goals. For more information, read the [Program Results Report](#).
- *Paths to Recovery*, 2002 to 2008, was designed to increase access to substance abuse treatment by improving the quality and efficiency of the delivery system at the provider level. Participating agencies used strategies originally developed by private industry that emphasized making a series of small, rapid-cycle changes to improve quality. For more information, read the [Program Results Report](#).
- *Transforming Care at the Bedside*, 2003 to 2008, developed, tested, and disseminated a structured process of continuous quality improvement for empowering nurses and other front-line hospital staff to take the lead in improving the work environment and the quality of patient care on medical-surgical units. For more information, read the [Program Results Report](#).

Evaluating these programs proved difficult, said RWJF Senior Program Officer Lori Melichar, PhD, MA, because the science was not at a level to answer confidently questions such as:

- How can we tell whether an initiative that uses the theory and methods of CQI to improve health outcomes has worked?
- How, and what, do we learn from these initiatives?
- How do we know whether a successful initiative will work in other organizations?

The core question was: “How will we know that the change is an improvement?”

Identifying Gaps in Research Methods

In 2007, RWJF engaged consultants Brian Mittman, PhD,² and Susanne Salem-Schatz, ScD,³ to interview directors, evaluators, and participants in RWJF national programs who were using CQI to try to improve health and health care, as well as leaders in the quality improvement field.

Mittman and Salem-Schatz found a number of barriers to using CQI in health care settings. For one thing, there is often a mismatch between conventional health care

¹ The PDSA cycle promotes continual improvement through planning (with attention to historical and current data), doing (pilot testing improvement interventions on a small scale), studying the pilot test results, and acting by implementing the improvements if the pilots are successful.

² Mittman is director of the VA Center for Implementation Practice and Research Support, Department of Veterans Affairs Greater Los Angeles Healthcare System.

³ Salem-Schatz is Principal, HealthCare Quality Initiatives, Newton, Mass.

clinical research methods and CQI work.⁴ CQI practitioners, researchers, and consumers also have few opportunities to examine peer-reviewed investigations of the success or failure of alternative CQI strategies.

Mittman and Salem-Schatz's recommendations for the field (see [Appendix 1](#)) helped RWJF set the research agenda for the *Improving the Science of Continuous Quality Improvement Program and Evaluation* program.

RWJF also created a “network,” through which the evaluators of RWJF quality improvement projects could confer regularly by phone, and produced summaries of the evaluations of eight RWJF quality improvement programs.⁵ The summaries shared a common template, “so that we could begin talking about quality improvement initiatives in a similar fashion,” Melichar said.

Funding

The RWJF Board of Trustees authorized *Improving the Science of Continuous Quality Improvement Program and Evaluation* for up to \$1.5 million for 48 months, from August 2007 through August 2011.

THE PROGRAM

Improving the Science of Continuous Quality Improvement Program and Evaluation aimed to address, in the health care environment, the core question: “How will we know that change is an improvement?” RWJF requested proposals in February 2008 to:

- Develop methods, tools, and standards for the design, conduct, and reporting of CQI research and evaluations
- Demonstrate how paradigms, methods, and approaches from other disciplines can advance the science of CQI research
- Develop and test new theoretical models, conceptual frameworks, and logic models to explain CQI initiatives
- Demonstrate how rapid-cycle “run-chart” or “PDSA-cycle” data can be used to rigorously evaluate the effectiveness of CQI interventions

⁴ For instance, health care clinical research methods do not support the use of logic models and theory-based hypothesis testing, do not focus on processes and impact, and do not acknowledge the importance of context and the adaptation of CQI programs to local circumstances. While disciplines such as education and the social sciences offer alternative evaluation approaches, these are not widely used in health care CQI. Also, the range of theories and models used for CQI research and evaluation is too narrow and not ideal for the study of health care delivery systems.

⁵ The RWJF programs were the *Diabetes Initiative*, *Expecting Success*, *Improving Chronic Illness Care*, *Paths to Recovery*, *Prescription for Health*, *Pursuing Perfection*, *Quality Allies*, and *Transforming Care at the Bedside*. Program Results Reports for all but *Quality Allies* are available at www.rwjf.org.

- Develop methods to systematically capture implementation context information
- Address other concepts that will advance the science of health care CQI research

Management

RWJF Research and Evaluation staff, led by Melichar, managed the program internally.

Grantee Meetings

Melichar and other RWJF Research and Evaluation staff convened two meetings of grantees and non-grantee experts in quality improvement at the Foundation’s offices in Princeton, N.J.:

- At the October 20, 2008 “kickoff” meeting participants discussed CQI definitions and the application of CQI in health care and grantees presented their intended projects.
- A March 9, 2011 meeting brought together grantees of this program and a related one, *Evaluating Quality Improvement Training Programs*, in a new community called Advance the Science of Quality Improvement Research and Evaluation (ASQUIRE⁶).

Funded Projects

From the proposals received, a panel of CQI researchers and other stakeholders recommended projects to RWJF staff, who selected these nine projects for funding. They fall into three groups

- Group 1: Creating a framework for the identification, classification, and evaluation of quality improvement initiatives (one project):
 - [RAND Corporation](#), Santa Monica, Calif. (ID#s 65113 and 67890)
- Group 2: Developing new quality improvement measures (five projects):
 - [Case Western Reserve University, Frances Payne Bolton School of Nursing](#), Cleveland (ID# 65114)
 - [Case Western Reserve University School of Medicine, Department of Medicine](#), Cleveland (ID# 65213)
 - [Cincinnati Children’s Hospital Medical Center](#), Cincinnati (ID# 65149)
 - [Johns Hopkins Bloomberg School of Public Health](#), Baltimore (ID# 65248)
 - [University of Medicine and Dentistry of New Jersey \(UMDNJ\) Medical School](#), New Brunswick, N.J. (ID# 65242)

⁶ For more information on this program, see [Evaluating Quality Improvement Training Programs](#) in this report.

- Group 3: Addressing data collection and analysis methodologies (three projects):
 - [Beth Israel Deaconess Medical Center](#), Boston (ID# 65121)
 - [Medical University of South Carolina](#), Charleston, S.C. (ID# 65188)
 - [Rutgers, the State University of New Jersey](#), Piscataway, N.J. (ID# 65181)

Most grants ran from October 2008 through December 2010. Awards varied from about \$50,000 to \$245,000. See [Appendix 2](#) for grant details.

PROJECT ACTIVITIES AND KEY RESULTS & FINDINGS

Group 1: Creating a framework to identify, classify, and evaluate quality improvement initiatives

Providing a Framework for Identifying, Classifying, and Evaluating Continuous Quality Improvement Studies⁷

This project, led by Lisa V. Rubenstein, MD, MSPH, at RAND Corp. aimed to develop evidence-based methods to identify and apply relevant quality criteria to published articles that empirically evaluate CQI literature. See the [Grantee Profile](#) of Rubenstein.

To do this Rubenstein and colleagues:

- Developed strategies to identify journal articles that evaluate CQI interventions. These included electronic search, title and abstract review, and full-text review.
- Engaged a panel of quality improvement stakeholders (government and university researchers, journal editors, and directors of organizations focused on quality improvement) to rate definitions and quality review criteria and to discuss preliminary results of the application of their ratings to journal articles identified through electronic searches.
 - They drew a list of definitions from language used in grantee proposals for the program,⁸ supplemented by some review of the literature, according to Rubenstein.
 - The [SQUIRE⁹](#) Guidelines were the initial basis for the quality review criteria, along with other lists of criteria that the project team identified.

⁷ ID#s 65113 and 67890.

⁸ This list was compiled by RWJF Research and Evaluation staff and presented at the initial grantee meeting on October 30, 2008.

⁹[SQUIRE](#) (Standards for Quality Improvement Reporting Excellence) Guidelines assist authors writing about quality improvement in healthcare to produce usable articles so that findings may be easily identified and widely disseminated.

- Identified and applied the quality criteria to a representative sample of articles that empirically evaluate CQI interventions, documented the level of adherence to these criteria, and analyzed the relationship between the criteria and each study.

The project team designed the Minimum Quality Criteria Set (MQCS) to assist in developing and evaluating scientific literature that:

- Reports empirical data on results of evaluations of quality improvement interventions
- Reports effects on patient health outcomes or clinical processes of care
- Aims to produce and test potentially enduring changes within the health care system, within community organizations, or within units within larger organizations

The stakeholder panel rated each of the 13 domains of the MQCS on whether the domain was appropriate for evaluating quality improvement publications. The domains included:

- Description of health-related outcomes
- Description of ability to be spread or replicated
- Description of organizational readiness for the studied intervention
- Initiated a large-scale, interactive, online panel process using the ExpertLens system developed at RAND for engaging large, diverse groups of stakeholders.
 - The RAND ExpertLens approach consisted of three one-week phases:
 - Week 1: Rating of 11 potential core CQI features¹⁰ using four measures¹¹
 - Week 2: Discussion of ratings
 - Week 3: Re-rating of CQI features and the addition of others suggested during the Week 2 discussion
 - Some 119 professionals participated. These included [Institute for Healthcare Improvement \(IHI\)](#) faculty,¹² editorial board members from quality improvement research journals, evaluators of RWJF quality programs, RAND patient safety and quality improvement experts, and other quality improvement professionals and health services researchers.

¹⁰ Examples of features: systematic data-guided activities; specific pre-identified aims; iterative development and testing; data feedback to implementers

¹¹ Measures: importance to the definition of a CQI initiative; importance to the success of a CQI initiative; importance of reporting the specifics in a publication; feature essential for review as a CQI initiative

¹² IHI is an independent nonprofit-organization based in Cambridge, Mass. It has managed a number of RWJF national programs, including *Pursuing Perfection*.

Key Results

Researchers reported key results on identifying CQI articles to RWJF:

- The CQI screen for identifying articles for key CQI definitional features developed through this project includes the following six items:
 - *Iterative Development Process*: Did the improvement initiative involve iterative design and implementation of a set of specific changes for improving care (i.e., a change package)?
 - *Feedback at Meetings Involving Participant Leaders*: Did leaders of the improvement initiative (e.g., local managers, clinical leaders, central experts, or improvement teams) from participating study organization(s) or local study site(s) meet to review information on initiative implementation?
 - *Feedback of Systematically-Collected Data*: Did the improvement initiative include feedback of systematically-collected data on implementation?
 - *Recognized Change Method*: Were one or more recognized change methods used in the improvement initiative?
 - *Data Driven*: In your judgment, to what extent was the design and/or implementation of a set of changes for improving care (i.e., change package) driven by data collected systematically during implementation?
 - *Local Conditions*: In your judgment, to what extent were local conditions at study organization(s) or site(s) taken into account in the design and/or implementation of the set of specific changes for improving care (i.e., the change package)?

Researchers reported on the application of the CQI screen to 106 articles the subject of which met the general definition of a quality improvement intervention in a 2011 article entitled “[Identifying Continuous Quality Improvement Publications: What Makes an Improvement Intervention ‘CQI’?](#)”¹³ published in *BMJ Quality & Safety*:

- The presence of the six features in the articles was as follows:
 - Feedback of systematically collected data: 64 percent of articles
 - At least “somewhat” adapted to local conditions: 61 percent of articles
 - Feedback at meetings involving participant leaders: 46 percent of articles
 - Using an iterative development process: 40 percent of articles
 - At least “somewhat” data driven: 34 percent of articles

¹³ O’Neill SM, Hempel S, Lim Y-W, Danz MS, Foy R, Suttorp MJ, Shekelle PG and Rubenstein LV. “Identifying Continuous Quality Improvement Publications: What Makes an Improvement Intervention ‘CQI’?” *BMJ Quality & Safety*, 20: 1011–1019, 2011. Available [online](#).

- Using a recognized change method: 28 percent of articles
- All six features were present in 14.2 percent of the articles.

Researchers reported on the experience of the ExpertLens process in a 2011 article entitled “[Conducting Online Expert Panels: A Feasibility and Experimental Replicability Study](#),”¹⁴ published in *BMC Medical Research Methodology*:

- Some 66 percent of participants completed all three phases. Of Phase I participants, 62 percent contributed to the discussions in Phase II, and 87 percent of those completed Phase III.
- Disagreement decreased after group discussion. After Phase III the experts agreed on the status of five of the 11 CQI features.
- The research team concluded, “It is feasible to conduct online expert panels intended to facilitate consensus finding among geographically distributed participants.”

Lessons About Identifying, Classifying, and Evaluating Continuous Quality Improvement Studies

In an interview, Project Director Rubenstein offered a set of lessons learned from the project.

- The literature is heterogeneous.
- Key definitions in the field are difficult and people do not agree on them.
- Literature searching in the field must be systematic. For the field to learn from its results, the literature searching capabilities need to be improved.
- People in the field need to think about hand screening after electronic searching.
- There are three communities of people engaged in this field and they are all going to have to come together at some point on these definitions:
 - Researchers
 - Practitioners
 - Journal editors/evidence-based practice center people who are focused on the literature
- There are tools available that can be adapted for advancing this field, like the ExpertLens method, that can help reach across different groups.

¹⁴ Khodyakov D, Hempel S, Rubenstein L, Shekelle P, Foy R, Salem-Schatz S, O’Neill S, Danz M and Dalal S. “Conducting Online Expert Panels: A Feasibility and Experimental Replicability Study.” *BMC Medical Research Methodology*, 11: 174, 2011. Available [online](#).

- Potentially machine learning or another systematic way can work with the difficulties in the wording in the field.
- It will take continued work in the field to get criteria that can be applied reliably.
- This foundational work has to be done if this field is going to add value in the long run.

Journal Articles

Danz MS, Rubenstein LV, Hempel S, Foy R, Suttorp M, Farmer MM and Shekelle PG. “Identifying Quality Improvement Intervention Evaluations: Is Consensus Achievable?” *Quality and Safety in Health Care*, 19(4): 279–283, 2010. Available [online](#).

Hempel S, Rubenstein LV, Shanman RM, Foy R, Golder S, Danz M and Shekelle PG. “Identifying Quality Improvement Intervention Publications—Comparison of Electronic Search Strategies.” *Implementation Science*, 6: 85, 2011. Available [online](#).

Khodyakov D, Hempel S, Rubenstein L, Shekelle P, Foy R, Salem-Schatz S, O’Neill S, Danz M and Dalal S. “Conducting Online Expert Panels: A Feasibility and Experimental Replicability Study.” *BMC Medical Research Methodology*, 11: 174, 2011. Available [online](#).

O’Neill SM, Hempel S, Lim Y-W, Danz MS, Foy R, Suttorp MJ, Shekelle PG and Rubenstein LV. “Identifying Continuous Quality Improvement Publications: What Makes an Improvement Intervention ‘CQI?’” *BMJ Quality & Safety*, 20: 1011–1019, 2011. Available [online](#).

Group 2: Projects to develop new quality improvement measures

Five projects sought to develop new measures that capture key inputs and outcomes of quality improvement efforts, as well as to define, measure, and report information about quality improvement implementation and its context.

A Measure of Systems Thinking: A Key Component of the Advancement of the Science of CQI¹⁵

Mary A. Dolansky, RN, PhD, and **Shirley M. Moore, RN, PhD**, researchers at Frances Payne Bolton School of Nursing at Case Western Reserve University developed and conducted a psychometric test¹⁶ of systems thinking.

Dolansky and Moore define systems thinking in this way: “The ability to recognize, understand, and synthesize the interactions, and interdependencies in a set of components designed for a specific purpose. This includes the ability to recognize patterns and

¹⁵ ID# 65114

¹⁶ A psychometric test measures a psychological aspect of a person, such as ability, motivation, or personality.

repetitions in the interactions and an understanding of how actions and components can reinforce or counteract each other.”

Systems thinking is considered to be critical to the success of CQI initiatives, yet no valid and reliable instrument existed to measure systems thinking. To meet this need, researchers:

- Explored the definition and elements of systems thinking through a literature search and online discussions of an expert panel
- Created a set of potential items for inclusion in the instrument and tested its content validity (i.e., the extent to which the instrument covered all aspects of systems thinking) with experts. They also field-tested the instrument to ensure that the meaning of each item was clear and that the format for responding was feasible.

The instrument asks respondents to indicate the extent (from “never” to “most of the time”) to which, when they want to make an improvement, they take a range of actions or consider a variety of points, such as:

- Seek everyone’s view of the situation
- Include people from their work unit to find a solution
- Think that systems are constantly changing
- Keep in mind that proposed changes can affect the whole system
- Recognize that system problems are influenced by past events
- Established discriminant validity (i.e., the extent to which items that should not be related are, in fact, not related) by testing with medical, nursing, and graduate health professions students who received varying degrees of training in systems thinking
- Performed other statistical tests to determine the overall reliability and validity of the instrument.

Key Results

Researcher reported key results to RWJF:

- **Researchers established the reliability and validity of the 20-item Systems Thinking Scale.** As a result, they said: “The first Systems Thinking Scale is available for use to assist in understanding one of the mechanisms by which CQI processes achieve their results—the ability to think about interdependencies in systems.”
- **Students who participated in the most intense training showed a significant increase in systems thinking, while those who received less or no training did not.**

- **Researchers also developed 17 additional items for two subscales (personal effort¹⁷ and reliance on authority¹⁸) for further testing.**
- **Researchers disseminated the Systems Thinking Scale to health care professionals involved in quality improvement projects and educators teaching systems thinking in health professions education through:**
 - A [website](#)¹⁹ that includes a description of the project, the Systems Thinking Scale instrument, an instrument manual, references, and other information
 - Presentations at the annual meeting of AcademyHealth (June 2009), at the International Health Forum on Quality and Safety in Healthcare (April 2010), and at the 16th Annual International Scientific Symposium on Improving the Quality and Value of Health Care (December 2010)

Journal Articles

Moore SM, Dolansky MA, Singh M, Caron A, Palmieri P and Alemi F. “Psychometric Analysis of the Systems Thinking Scale.” Unpublished.

Presentations

Moore SM and Dolansky MA. “The Psychometric Testing of an Instrument to Measure Systems Thinking,” at the International Improvement Science Symposium; April 2012, Paris.

***Assessing the Implementation Context: A Mixed Methods Approach*²⁰**

David Litaker, MD, PhD, and colleagues in the Division of Family Medicine Research at Case Western Reserve University sought to produce a reliable and valid instrument to measure key features of primary care practices that influence the adoption and implementation of new knowledge, such as evidence-based care.

While much scientific evidence to guide health care delivery is available, there are many gaps in the quality of care delivered in primary care—and one factor is the differences in practices’ capacity for change, according to Litaker and his research team.²¹ To address this issue, researchers:

¹⁷ Examples of personal effort items include agreement with: “Lasting change relies on personal effort and motivation” and “It is not how hard you work, but how smart you work.”

¹⁸ Examples of reliance on authority items include agreement with: “The leaders of the organization have the best ideas” and “A top-down strategy is usually best.”

¹⁹ [The Systems Thinking Scale: A Measure of Systems Thinking.](#)

²⁰ ID# 65213

²¹ Litaker D, Ruhe M and Flocke S. “Making Sense of Primary Care Practices’ Capacity for Change.” *Translational Research*, 152: 245–253, 2008.

- Created a set of 117 items related to practice capacity for change, based on review of previous conceptual and empirical studies
- Rated these items, using a five-point Likert scale (strongly disagree to strongly agree), using data (direct observation and key informant interviews) from 15 primary care practices engaged in a quality improvement intervention
- Conducted a range of statistical tests and modeling techniques to reduce the number of items and finalize the instrument
- Assessed the instrument’s reliability²² and convergent validity²³

Key Results

Researchers reported key results in a 2009 article entitled “Measuring Practice Capacity for Change: A Tool for Guiding Quality Improvement in Primary Care Settings,”²⁴ published in *Quality Management in Health Care*.

- **The final 25-item instrument had strong reliability.** The instrument also had good convergent validity.
- **Some of the 25 final items included in the instrument are:**
 - There is a lot of tension between people in this practice.
 - This practice has a shared vision among practice members.
 - This practice generally sees change as possible.
 - This practice is able to stay on schedule.
 - Things change so fast in this practice it’s hard to keep up with what’s going on.
- **The project team concluded that “the ability to quantify capacity for change may enable better recognition of practices likely to be successful in their change efforts and those first requiring capacity building prior to change interventions.”**

Journal Articles

Bobiak SN, Zyzanski SJ, Ruhe MC et al. “Measuring Practice Capacity for Change: A Tool for Guiding Quality Improvement in Primary Care Settings.” *Quality Management in Health Care*, 18(4): 278–284, 2009. Abstract available [online](#). (Scroll down and click on article.)

²² Reliability: the consistency of a test or measure.

²³ Convergent validity: whether items that should be related to each other are actually observed to be related.

²⁴ Bobiak SN, Zyzanski SJ, Ruhe MC, et al. “Measuring Practice Capacity for Change: A Tool for Guiding Quality Improvement in Primary Care Settings.” *Quality Management in Health Care*, 18(4): 278–284, 2009. Abstract available [online](#). (Scroll down and click on article.)

Ruhe MC, Carter C, Litaker D and Stange KC. “A Systematic Approach to Practice Assessment and Quality Improvement Intervention Tailoring.” *Quality Management in Health Care*, 18(4): 268–277, 2009. Abstract available [online](#). (Scroll down and click on article.)

Litaker D, Ruhe M, Stange K and Gao Y. “Development of a New Tool to Assess Primary Care Practices’ Ability to Translate New Knowledge into Action.” Unpublished.

Developing a Framework for Optimizing Contextual Factors That Impact Quality Improvement Projects and Their Likelihood of Success

Heather C. Kaplan, MD, MSCE, and colleagues at Cincinnati Children’s Hospital Medical Center developed a framework that organizations and implementation researchers can use to study, understand, and optimize contextual factors that impact a quality improvement project’s likelihood of success. See the [Grantee Profile](#) of Kaplan.

“There is tremendous variation in how successful quality improvement projects are,” says Kaplan. “We believe that QI methods work—there is evidence from other industries that they work and a long history that they work. So if QI methods are effective, then the variation has to be related to the context.”

To develop and test a framework for optimizing contextual factors in quality improvement Kaplan and her team:

- Reviewed both the health care and the business literature, considering over 13,000 titles and abstracting data on nearly 100 articles
- Engaged 10 quality improvement experts in the model development process that included consideration of the literature review results and two rounds of opinion gathering to identify important contextual factors and clarify definitions. In an in-person meeting, panelists identified relationships among the factors.
- Tested the resulting Model for Understanding Success in Quality (MUSIQ) in 74 quality improvement projects operating in three different settings through a web-based questionnaire completed by staff from these projects:
 - Cincinnati Children’s Hospital Medical Center’s Strategic Improvement Priority Initiatives (43 projects)
 - Ohio Perinatal Quality Collaborative obstetrical and neonatal improvement projects (19 projects)
 - Institute for Healthcare Improvement’s Improvement Advisor Program (12 projects)

Key Findings

Researchers reported key findings from the 47 health care articles included in the final literature review in a 2010 article, “[The Influence of Context on Quality Improvement Success in Health Care: A Systematic Review of the Literature](#),” published in *Milbank Quarterly*.²⁵

- **Factors suggested as important to the success of quality improvement projects include:**
 - Leadership from top management
 - Organizational culture
 - Data infrastructure and information systems
 - Years involved in quality improvement
- **Other potentially important factors include:**
 - Physician involvement in quality improvement
 - Micro-system (i.e., unit/department/office level) motivation to change
 - Resources for quality improvement
 - Quality improvement team leadership
- **Researchers noted that the existing literature lacked:**
 - A practical conceptual model
 - Clear definitions of contextual factors
 - Well-specified measures

Results of the systematic review along with input from an expert panel were used to develop the Model for Understanding Success in Quality (MUSIQ). Researchers described key elements of MUSIQ in an article entitled “[The Model for Understanding Success in Quality \(MUSIQ\): Building a Theory of Context in Healthcare Quality Improvement](#),” published in 2012 in *BMJ Quality & Safety*.²⁶

²⁵ Kaplan HC, Brady PW, Dritz MC, Hooper DK, Linaam WM, Froehle CM and Margolis P. “The Influence of Context on Quality Improvement Success in Health Care: A Systematic Review of the Literature.” *Milbank Quarterly*, 88(4): 500–559, 2010. Abstract available [online](#).

²⁶ Kaplan HC, Provost LP, Froehle CM and Margolis PA. “The Model for Understanding Success in Quality (MUSIQ): Building a Theory of Context in Healthcare Quality Improvement.” *BMJ Quality & Safety*, 21: 13–20, 2012. Abstract available [online](#).

- **MUSIQ identifies 25 key contextual factors, at different levels of the health care system, likely to influence quality improvement success.** Examples at each level include:
 - *External environment*: CQI project sponsorship
 - *Organization*: Senior leaders’ commitment to champion the project; a culture supportive of quality improvement
 - *Quality improvement support and capacity*: Data infrastructure; resource availability
 - *Micro-system*: Capacity for improvement; motivation for change
 - *Quality Improvement team*: Team diversity; physician involvement; team leadership
 - *Miscellaneous*: Presence of a trigger event (positive or negative) that stimulates emphasis on improving quality
- **MUSIQ also hypothesizes relationships among contextual factors and success in quality improvement:**
 - Micro-system (i.e., unit/department/office level) and quality improvement team factors are hypothesized to *directly* affect quality improvement success.
 - Factors in the external environment and at the organization level are hypothesized to influence quality improvement success *indirectly*.

Researchers reported preliminary results of the model testing in 74 QI projects to RWJF:

- **Contextual factors with the greatest total effects on quality improvement success include:**
 - Resource availability
 - Quality improvement team leadership, team quality improvement skills, and quality improvement team decision-making processes
 - Micro-system motivation and quality improvement capability
- **Other contextual factors that may have important effects on quality improvement success among the participating projects include:**
 - Organizational quality improvement culture and senior leader project sponsorship
 - Micro-system quality improvement leadership and quality improvement culture

“This two-phase project worked well,” said Kaplan, “with key learnings in the first phase that informed the second. The literature review supported our initial hunch that context was important in QI and that it was essential to start to attack it in a systematic way. Then

the testing we did with the quality improvement projects gave us the confidence that we had established more validity for our model and were on the right track. Now we can really delve in and advance it further.”

Journal Articles

Kaplan HC, Brady PW, Dritz MC, Hooper DK, Linaam WM, Froehle CM and Margolis P. “The Influence of Context on Quality Improvement Success in Health Care: A Systematic Review of the Literature.” *Milbank Quarterly*, 88(4): 500–559, 2010. Abstract available [online](#). Full text requires subscription or fee.

Kaplan HC, Provost LP, Froehle CM and Margolis PA. “The Model for Understanding Success in Quality (MUSIQ): Building a Theory of Context in Healthcare Quality Improvement.” *BMJ Quality & Safety*, 21: 13–20, 2012. Abstract available [online](#).

Capturing Implementation Context: The Team Checkup Tool²⁷

This project, led by **Jill A. Marsteller, PhD, MPP**, at Johns Hopkins Bloomberg School of Public Health, aimed to inform understanding of the measurement and role of implementation context in continuous quality improvement initiatives by validating the Team Check-up Tool, improving it, and making it more generalizable.

The Team Check-up Tool had been used in ICUs in two collaboratives aimed at reducing bloodstream infection—in the intensive care units at Johns Hopkins Hospital and in 103 ICUs participating in the Michigan Keystone Intensive Care Unit Project.²⁸ Infections were greatly reduced, but why the intervention worked was not known. To understand variation in outcomes of team-based CQI efforts, the Johns Hopkins team believed the traits and activities of the quality improvement teams themselves must be considered.

The tool measures health care CQI team characteristics and activities. Its purpose is to help project, hospital, and team leaders work with teams to resolve problems that prevent goal achievement. Completed by the team leader monthly or quarterly, the Tool provides information about team activities and problems that slowed progress during the previous month or quarter. Among the items included:

- Nurse, physician, and executive leadership and buy-in related to the intervention being implemented
- Performance review and communication with staff regarding the intervention and results
- Knowledge of the evidence base for the intervention

²⁷ ID# 65248

²⁸ The Michigan Keystone Intensive Care Unit Project was a partnership between the Michigan Health & Hospital Association and Johns Hopkins University.

- Availability of time

It also assesses the quality improvement team’s “inability to work together” and interpersonal problems a team might face, such as:

- Insufficient participation by one or more team members
- Some members not valuing the contributions of other members
- Personality conflicts

In this project Marsteller and colleagues evaluated the Team Check-up Tool in its use by quality improvement teams working with 46 ICUs in 35 faith-based community hospitals in 12 states for the following:

- Face validity (whether a test seems to measure what it is intended to measure) and feasibility, using focus groups of team members from three collaboratives to reduce central line bloodstream infections. Researchers hoped that focus group information could be used to support use of the tool in a range of settings and CQI contexts beyond the ICU, for which it had been originally developed.
- Content validity (whether a measure addresses all aspects of the concept it is designed to measure), using an expert panel and a survey
- Reliability (the consistency of a test or measure) and variability (how spread out or closely clustered a set of data is), through analysis of existing data from a Johns Hopkins CQI program
- Construct validity (whether what was to be measured was actually measured), convergent (whether items that should be related to each other are actually observed to be related) and discriminant (whether items that should not be related to each other are actually observed to be not related) validity, and predictive validity (the extent to which scores on one test are predictive of scores on another), using secondary data analysis.

Key Findings

Researchers reported key findings in a 2011 article “[Validity and Usefulness of Members Reports of Implementation Progress in a Quality Improvement Initiative: Findings from the Team Check-up Tool \(TCT\)](#),” published in *Implementation Science*²⁹:

- The Team Check-up Tool showed good measurement reliability, validity, and responsiveness.

²⁹ Chan KS, Hsu Y-J, Lubomski L, and Marsteller JA. “Validity and Usefulness of Members Reports of Implementation Progress in a Quality Improvement Initiative: Findings from the Team Check-up Tool (TCT).” *Implementation Science*, 6: 115, 2011. Available [online](#).

- Researchers concluded that with “more validated measures on implementation context, researchers can more readily conduct rigorous studies to identify contextual variables linked to key intervention and patient outcomes and strengthen the evidence base on successful spread of efficacious team-based interventions.”

Key Results

- Marsteller reported to RWJF that the Team Check-up Tool has been used by ICUs in more than 35 states that are participating in a national project to reduce bloodstream infections (STOP BSI) funded by the Agency for Healthcare Research and Quality (AHRQ). Other types of CQI initiatives also have used the tool, including a project that aimed to improve congestive heart failure in rural hospitals and another, also funded by AHRQ, on cardiac surgical treatment in 17 hospitals.

Journal Article

Chan KS, Hsu Y-J, Lubomski LH and Marsteller JA. “Validity and Usefulness of Members Reports of Implementation Progress in a Quality Improvement Initiative: Findings from the Team Check-up Tool (TCT).” *Implementation Science*, 6: 115, 2011. Available [online](#).

Enhancing the Theory and Practice of Continuous Quality Improvement through a Better Understanding of Organizational Capacities³⁰

Deborah Cohen, PhD, Elizabeth C. Clark, MD, MPH, and colleagues at the University of Medicine and Dentistry of New Jersey addressed the question: “Why are some primary care practices able to integrate and sustain CQI efforts while others are not?”

The research team developed a theory to expand understanding of the factors that enhance or inhibit the ability of primary care practices to integrate and sustain CQI efforts. The model emerged from analysis of audio recordings of quality improvement team meetings, which provided insight into the dynamics of the quality improvement process, and from analysis of quantitative data. The primary care practices involved had participated in one of four quality improvement studies.

The model identifies factors that influence the engagement (i.e., active participation through dedicated time and action) of a primary care practice in quality improvement work.

Key Findings

In a presentation entitled “Development of a Theory for Understanding Engagement in Quality Improvement in Primary Care”³¹ at the Annual Research Meeting of the North

³⁰ ID# 65242

American Primary Care Research Group in 2010, Cohen and Clark reported on the key factors in the model:

- Professional status—the relative position of individuals within the practice
- Psychological safety—the sense of safety individuals have within a group. Individuals are more likely to be active participants when they feel psychologically safe in the group.
- Leadership behavior—the actions of those in charge that may influence the involvement of others.
- Whether the focus of the quality improvement effort was clinical or administrative. This defines who should be involved in the quality improvement process.

In summary, Cohen and Clark suggested that “practices able to manage differences in professional status create psychologically safe environments in which quality improvement engagement is more likely to result.”

Group 3: Projects to address data collection and analysis methodologies

Determining the Reliability of a Common Tool Used to Measure Continuous Quality Improvement³²

This project, led by **Michael D. Howell, MD, MPH**, at Beth Israel Deaconess Medical Center, addressed the concern that quality improvement studies that express outcomes as “events per risk-exposure-time” (such as “falls per 1,000 patient-days”) may inaccurately reflect actual outcomes over time. See [Grantee Profile](#) of Howell.

Howell’s team sought to develop more reliable metrics for use in future research, thereby enhancing the quality and generalizability of findings from common targets of quality improvement research (such as fall reduction and ventilator-associated pneumonia prevention).

Researchers investigated the reliability of a very common metric: “falls per 1,000 patient-days,” that is used by quality improvement teams, payers, and regulatory agencies that compare hospitals. “Falls were attractive for three reasons,” says Howell, “they affect a lot people around the country, there is a lot of interest in them, and we could get data that would let us test the hypothesis.” The researchers:

- Created and built a model of the intersection of risks that vary over time and the variation in exposure-time

³¹Elizabeth C. Clark, “Development of a Theory for Understanding Engagement in Quality Improvement in Primary Care,” at the Annual Research Meeting of the North American Primary Care Research Group, November 16, 2010, Seattle. Abstract available [online](#). Scroll to bottom and enter presentation title.

³² ID# 65121

- Integrated data from the Patient Safety Reporting System and case mix (i.e., the mix of types of patients, diagnoses, etc.) systems at Beth Israel Deaconess into a dataset of more than 1,500 falls linked to the day of the fall (in the patient’s hospital stay). They drew this dataset from 121,865 discharges that accounted for 583,786 patient-days.
- Studied national empirically observed variations³³ in exposure time in the fall metric over a five-year period (2001 to 2006) for:
 - Changes over time for individual hospitals
 - Differences among hospitals in a single year
- Worked to develop a standard metric

Key Findings

Howell and his team reported their key findings to RWJF:

Fall Risk at the Individual Hospital Over Time

- **Fall risk at Beth Israel Deaconess varied across patient length of stay, with twice the risk of falling on day 14 as on day 1.** After day 14 (only a very small percentage of patients at this hospital had long lengths of stay) the risk of fall varied greatly from day to day.
- **From the national dataset researchers found that many individual hospitals had substantial variation in length of stay over a five-year period.** Of the 244 hospitals present in both the 2001 and 2006 samples, 26 percent had a change in median length of stay of one day or more during that time.
- **Using median length of stay substantially underestimated variability among hospitals.** Hospitals with very different distributions in lengths of stay (e.g., from a wide range of stays to having most stays cluster around the median) could have the same or very similar median lengths of stay. Using the median does not uncover many of the differences among hospitals.
- **Using a simulation model in which fall risk was identical in each hospital from year to year, a substantial number of hospitals would have erroneously concluded that they had either improved or worsened by a clinically meaningful amount.** These conclusions would have been incorrect.
 - Some 29 percent would have concluded that their fall performance improved or worsened by 10 percent or more.

³³ From the National Inpatient Sample of the Agency for Healthcare Research and Quality’s [Healthcare Cost and Utilization Project](#), which includes the largest collection of longitudinal hospital care data in the United States, from 1988.

- Some 10 percent would have concluded that performance changed by at least 20 percent.
- Some 5 percent would have seen at least a 30 percent variation in fall risk.
- However, the researchers noted that random variation in sampling was also a significant factor in this distortion.

Fall Risk Across Hospitals

- **From the national dataset, researchers found that length of stay varied widely among hospitals.** While median length of stay was three days, 5 percent of hospitals had a median length of stay of more than six days. And, as noted above, hospitals with the same median length of stay could have very different distributions in stays.
- **Using a simulation model in which fall risk was identical in all hospitals, researchers found that substantial percentages of hospitals would be erroneously found to differ from each other.** This variation was mostly driven by the differences in length of stay distributions. In this model, despite an identical fall risk in all hospitals:
 - Some 59 percent of hospitals differed from the national median “falls per 1,000 patient-days” by 10 percent or more.
 - Some 30 percent varied from the national median by 20 percent or more.
 - Some 16 percent varied from the national median by 30 percent or more.

Key Challenges

Researchers reported two important challenges to RWJF:

- The observed (Beth Israel Deaconess) data included very few patients with long lengths of stay, so precision decreased as length of stay increased, due to small sample numbers. The researchers are not aware of a national dataset that includes hospital day on which the fall occurred linked to the patient’s length of stay. This would have been helpful to this analysis.
- Random chance had a big effect on individual hospitals’ results, especially for smaller hospitals with fewer patients.

Key Conclusions

Researchers offered the following conclusions in a report to RWJF:

- “The current standard metric for a very common CQI target (inpatient falls) is fundamentally flawed and will frequently result in erroneous conclusions by quality improvement teams and policy-makers.”

Project Director Howell concluded in an interview:

- “The most important result of this study is that subtle things in quality measurement can really, really matter. Naïve denominators [i.e., those that do not account for real, underlying variation where, for example, one patient day is not the same as another in terms of risk] can make the metrics so flawed as to give the wrong answer, both at a hospital level and at a policy-maker level. The metrics implicitly assume that every patient day has the same risk as every other patient day. But it’s not the case. The implicit assumptions of the metric—which are basically an averaging problem—can lead to the wrong conclusion.”

Journal Article

Howell MD. “A 37-Year-Old Man Trying to Choose a High-Quality Hospital: Review of Hospital Quality Indicators.” *Journal of the American Medical Association*, 302(21): 2353–2360, 2009. Available [online](#).³⁴

Data Analysis and Inference Issues in Continuous Quality Improvement³⁵

Researchers in the Department of Psychiatry and Behavioral Sciences at Medical University of South Carolina, led by **Jeffrey J. Borckardt, PhD**, investigated the short-comings of industrial and manufacturing quality assurance techniques (called statistical process control) when applied to health care data. Previous research had shown that the human-generated data encountered in the health care setting (rather than the industrial machine data of the manufacturing environment) can violate fundamental statistical assumptions inherent in the industry-based data-analytic techniques, leading to errors in data interpretation.

For example, the industry-based techniques assume that if you collect enough data, it will fall in a normal distribution (bell-shaped curve)—and with machine-generated data that is likely to be the case. Human-generated data is much less likely to follow such a distribution.

Borckardt and colleagues noted the need for new, user-friendly statistical process control techniques for the health care setting. They:

- Conducted a detailed review of statistical process control data analytic techniques
- Identified their short-comings when applied to health care data, using Monte-Carlo statistical modeling techniques for simulations that rely on repeated random sampling

³⁴While Howell considers this article to come from this work, it does not report on this work, but was an article he was invited to write by an editor at *JAMA*.

³⁵ ID# 65188

- Developed a set of tools that are better suited to the unique properties of health care quality data

Key Findings

Borckardt reported key findings to RWJF:

- Violations of statistical assumptions associated with statistical process control techniques are common among health care quality data.
- Violations of the statistical assumption of “data normality” had little impact on false positive (i.e., a positive result when, in fact, the real result is negative) and false negative (i.e., a negative result when, in fact, the real result is positive) rates when the traditional statistical process control techniques were applied to health care data. Data normality refers to whether the data studied conform to a “normal” distribution (i.e., the bell-shaped curve) in which data cluster around a single mean value.
- Violations of the statistical assumption of “data independence” had a substantial impact on false positive rates but not on false negative rates. Data independence is the separation of the data from the computer programs that use the data (i.e., the data are not affected by changes in the program).
- The investigators developed and tested a data analytic technique that showed good control of false positive and false negative rates when assumptions of data normality and independence are violated.

Designing and Implementing Statistical Methods for Detecting Changes in Organizational Performance³⁶

Thomas O. Boucher, PhD, MBA, MS, and colleagues at Rutgers University, designed and implemented a software tool to help health care organizations detect changes in organizational performance during the introduction of new technology or when changes are made to a current process or work flow. This tool also estimates the new steady state level of performance in place after organizational learning has occurred.

The practical significance of this software is that it will provide a tool to automatically monitor quality-related measures for changes, and integrate the detection of learning behavior for forecasting final levels of improvement.

The research team tested and validated the software through use of actual data from research articles and from nonpublic proprietary data sets, and from computer simulations of data sets with known properties.

³⁶ ID# 65181

Key Results

Boucher and colleagues developed a two-part series of computer programs, with associated user's manuals, to provide statistical tools to support CQI projects:

- *The Improvement Analyzer: Automated Improvement Curve Detection and Modeling.* This program provides decision support for CQI team members who have implemented complex changes that require substantive learning by employees, such as a new computer program with new procedures and a need for personnel to adapt to new technology and processes.

In a complex change process, initial performance may be worse than before the change. It is important to be able to estimate, from early data, the final performance level after staff has mastered the new process or system. The Improvement Analyzer provides a tool to analyze the early data and use it to estimate the future performance level. It is designed to be used with manual data entry.

- *The Real-Time Performance Monitor: Real-Time Process Change Detection and Modeling.* This program provides real-time tracking of statistical changes in process measures that are being monitored by CQI project team members. It can be used on-line (to use data from a database) or off-line (for inputting data from an Excel spreadsheet). The program monitors a stream of data from a CQI project and detects changes from user-defined process metrics. In this way, the program can identify a change in learning or the beginning of a transition to a new level. It provides an early signal of a process change and whether the change is a process improvement.

Software and Manuals

The Improvement Analyzer: Automated Improvement Curve Detection and Modeling. Software and User's Manual. Piscataway, N.J.: Industrial & Systems Engineering, Rutgers University, 2011. User's manual available [online](#).

The Real-Time Performance Monitor: Real-Time Process Change Detection and Modeling. Software and User's Manual. Piscataway, N.J.: Industrial & Systems Engineering, Rutgers University, 2011. User's manual available [online](#).

LESSONS LEARNED FROM FUNDED PROJECTS

Lesson on Developing Definitions for CQI

1. **Use concrete research examples of current publications to address definitional issues related to CQI.** Relying only on theory is not sufficient for identifying deficiencies in definitions of CQI. (Project Director Lisa V. Rubenstein, RAND Corporation)

Lessons on Creating Quality Improvement Measures

- 2. When advancing the science in a field, include instrument development as a necessary step.** It is a key component of any field. (Project Director Mary A. Dolansky, Case Western Reserve University Frances Payne Bolton School of Nursing)
- 3. Tie measure development to conceptual models.** Developing measures to advance quality improvement research is challenging, especially validating measures of quality improvement success. A conceptual model, like Cincinnati Children's MUSIQ model, offers a way to think through the critical data needs clearly, and is a strong base on which to build a new measure. (Project Director Heather Kaplan, Cincinnati Children's Hospital)
- 4. Involve top experts in a field when doing basic research.** Their input is valuable. In particular, a psychometrician skilled in contemporary measurement strategies should be part of the research team. (Project Director Dolansky, Case Western Reserve University Frances Payne Bolton School of Nursing)
- 5. Consider a mixed methods approach when developing a new instrument.** Project Director David G. Litaker, at Case Western Reserve University School of Medicine, said that he was "very fortunate to work within a skilled, interdisciplinary group of researchers who were knowledgeable in the application and integration of both quantitative and qualitative techniques." He found the mixed methods approach to be "highly valuable in instrument development and refinement."
- 6. Strike a balance between obtaining sufficient detail through long, multi-item questionnaires and shortening questionnaires to limit respondent burden.** While respondents to the Cincinnati Children's Hospital survey reported that the survey was not too burdensome, "this approach [a short questionnaire] limited the depth of understanding we could achieve," said Project Director Kaplan. "There is likely a middle ground where the questionnaire could include more multi-item scales, while still limiting respondent burden."
- 7. Use personal outreach to boost recruitment of survey respondents.** Researchers at Cincinnati Children's Hospital tested various strategies to increase response rates and found personal outreach to have the greatest effect. (Project Director Kaplan, Cincinnati Children's Hospital)
- 8. Discuss early results with others in the field and engage well-respected experts in the model building process.** "Success of this project hinged on the community of researchers we built throughout the execution of the project," said Cincinnati Children's Hospital Project Director Kaplan. Testing a model at a project level requires collaborations with other researchers as well as quality improvement projects that can serve as laboratories for model testing.

9. **Continually improve an instrument after release of the first version.** With use in the field, changes needed to keep the instrument current and to refine its performance will become apparent. Researchers at Case Western Reserve University Frances Payne Bolton School of Nursing added additional items for two subscales to the original Systems Thinking Scale as they gained experience with the instrument. (Project Director Dolansky, Case Western Reserve University Frances Payne Bolton School of Nursing)

Lessons on Data Collection and Analysis Methodologies

10. **Be aware that accounting for sampling variation when working with very large data sets may be quite challenging.** It is important to consider the possible impact of this variation in designing a study and in interpreting findings. (Project Director Michael D. Howell, Beth Israel Deaconess Medical Center)
11. **When developing mathematical tools, collaborate with other researchers who have access to data and ongoing CQI project case studies that can be used to test the models and associated software products.** To ensure the usefulness of computer programs in practice, empirical testing is critical. (Project Director Thomas O. Boucher, Rutgers University)
12. **Consider using audio-recordings of quality team meetings as a data source for quality improvement studies.** By listening to and analyzing such audio-recordings, Project Director Elizabeth C. Clark at University of Medicine and Dentistry of New Jersey and her colleagues discovered many insights into the projects in their sample that otherwise would not have been available to them.
13. **Engage organizational leaders when planning data collection in order to get their buy-in, permission, and support.** Leaders of the two organizations being studied allowed researchers from Johns Hopkins Bloomberg School of Public Health to attend their organizational meetings and provided space for conducting focus groups. (Project Director Jill A. Marsteller, Johns Hopkins)
14. **Plan to use multiple data sources to allow for modifications of key objectives in response to changes “on the ground.”** Project Director Marsteller at Johns Hopkins Bloomberg School of Public Health said her team learned the value of this approach when they could not obtain one data set. They consequently pursued an additional data collection process that ended up strengthening the overall project.
15. **Be realistic about what can be achieved with available person-power and funding.** At Johns Hopkins Bloomberg School of Public Health, said Project Director Marsteller, this “allowed a more reasonable pace in the conduct of this study than I have seen in some other studies, and thus allowed us to think creatively about solutions, additions and new ideas.”

RELATED WORK

During the period of the CQI national program, RWJF funded several other projects and initiatives related to quality improvement.

Quality Improvement Training

Scanning the Field

From 2007 through 2011, researchers at the Boston University School of Management Health Policy Institute, led by Sally K. Holmes, MBA, supported efforts by RWJF's Human Capital Team to increase the number of health care workers who are highly proficient in quality improvement methods and tools.³⁷ To achieve this, researchers:

- Surveyed the field of quality improvement training to ascertain its present scope and availability, as well as to determine factors hindering or encouraging the use of training
- Developed online resources and programmatic options in quality improvement training for future deployment by the Human Capital Team

Researchers determined that, although numerous quality improvement training options exist, learners are not well connected with these resources. They also found that training in quality improvement methods and tools does not necessarily assure long-term adoption by staff members. The researchers concluded that more effective incentives are needed to foster widespread adoption of quality improvement in health care.

The results of this effort led to the development of a new RWJF national program to assess quality improvement training.

Evaluating Quality Improvement Training Programs

In 2008, RWJF launched a new, \$2 million national program, *Evaluating Quality Improvement Training Programs*. The program runs through 2012. It is internally managed by Foundation staff, and led by Senior Program Officer Lori Melichar, PhD, who also directed the CQI national program. The goal of this program is to increase the understanding of what works in quality improvement training programs so that more organizations adopt best practices and more health providers acquire training in quality improvement.

Program grantees are evaluating five different quality improvement training programs.

³⁷ See [Program Results Report](#) for grant ID# 65102.

A Community of Quality Improvement Researchers

Bringing Quality Improvement Researchers Together

RWJF's Melichar, working with consultant Holmes who led the scan of the quality improvement training field, brought the grantees of both national programs together in a meeting at RWJF in March 2011. Melichar said that the meeting's purpose was "to see how they might use common measures, to see what they thought would be their challenges, to try to really create a community that would be supportive to each other as they were conducting their projects."

An Online Community

In addition, Melichar and Holmes are developing (through May 2013) an online community as a venue for ongoing connection among the grantees of the two quality improvement-focused programs. Called ASQIRE: *Advancing Science of Quality Improvement Research and Evaluation*, the site has not caught on among the researchers as intended. Said Holmes in an interview in January 2012: "I would like to use the website as a community tool where we can facilitate interactions between grantees and other people who are key in the field. Right now it's not a very dynamic or interesting website and we need to improve that."

Melichar hopes that eventually ASQIRE will become a website with publicly available content and be "the go-to source for quality improvement research and evaluation." But, she says, for that to happen, "we will need to evaluate it and make more effort to engage people with it."

BUILDING A FIELD

Grantees and consultants involved in RWJF's quality improvement-focused programs and projects offered perspectives on building the science of quality improvement research, evaluation, and training:

The Field Is Broad and Evolving

- "The area of quality improvement evaluation, research, and training is a huge field. To get the impact that the Foundation is usually going for with its investments, you have to think very carefully about what piece of that field to grab hold of and go deep with, to really understand and say something definitive about. And, if you're able to do that, to figure out how to influence the thousands of other people out there who are working in this huge field." (Consultant Sally K. Holmes)
- "The biggest challenge is that the whole field of quality improvement is extraordinarily broad in health care and is also evolving very quickly. There is not

even a shared vocabulary at this point in time—though Rubenstein’s work at RAND is advancing that.” (Holmes)

- “This work really opened my eyes to how much progress needs to be made in this field, and that there is still a lot to be done. It’s really going to be a community effort of everyone involved in quality improvement to try to move the work forward.” (Cincinnati Children’s Hospital Project Director Kaplan)

The Field Requires New Standards and New Ways of Learning

- “It was difficult, almost impossible, to forge agreement [in a discussion at the March 2011 meeting] about using standard measures. Researchers acknowledged the desirability of doing that. But in their studies they had gone their different ways in focusing on the research questions and were not terribly interested in compromising their studies, as they envisioned them, by trying to adopt standard measures.

“In hindsight, perhaps we should have put more emphasis up front on the idea of shared measures. We thought that those would fall out of the research and that people would collaborate. If you want grantees to use standard measures, you have to set it up differently.” (Consultant Holmes)

- “There are times when it is reasonable to do randomized trials of a social-contextual intervention. But people in general do them way too early, when they are still trying to understand what the program is. People are struggling with what’s the right way to do this and trying to increase the acceptance of alternative methods of learning and not to be held hostage to the clinical model for learning.” (Consultant Susanne Salem-Schatz³⁸)
- In order to do these programs successfully, you need to adapt them to your own circumstance. It is really hard to know what is the ‘unmessable’ core—the things that you really have to include in order to call it this kind of intervention. CQI needs a different approach to learning and understanding.” (Salem-Schatz)
- “One of the challenges in this field is that there was such a lack of knowledge and research. It’s really starting to pick up now, but to be most helpful we need to do work that builds on top of each other and not just be duplicative. Communicating as early as possible in a safe, non-punitive, collaborative environment is important so that people aren’t just reinventing the wheel.” (Kaplan)
- “People who are trying to do quality improvement on the front lines really want practical tools to identify the risk of failure and why an intervention fails, and they want the knowledge to fix the problem. The research is lagging behind what the people on the front lines really want.” (Kaplan)

³⁸ See [Identifying Gaps in Research Methods](#).

Context Is Important

- “Everybody at this point agrees that context matters. Some people are particularly traditional in their approach to context and try to control for it. Others of us think it’s all about the context and how you make these changes and how you adapt for context. The last thing you want to do is to make that go away and say that, all other things being equal, this works. It’s about finding the right fit or approach for the given site and context.” (Salem-Schatz)

The Field Has Advanced Since the Inception of RWJF’s CQI Program

- “The field of quality improvement on the scientific end—whether you’re a practitioner or a researcher—lagged for a long time and it was kind of weird if you did it. Then all of a sudden the demand really went up.” (RAND Project Director Lisa V. Rubenstein)
- “What is encouraging in the field—and RWJF was probably part of setting the stage for this—is that people have a much clearer sense now about how career issues impact the ability to do this work. To this point, if you were a young investigator you would likely not choose this area to work in because you would not be likely to get easy funding and nor get published in big, important journals.” (Salem-Schatz)

The RWJF Role Has Been Critical in Advancing the Field

- “There is a lot of talk about the way we do quality improvement but not a lot of organized initiatives about how we actually do research about quality improvement. RWJF really played a role that nobody else was playing. I think we still remain the leader in that.” (RWJF Senior Program Officer Lori Melichar)
- “RWJF was on the cusp of funding this kind of work before it really started to take off. This program was very helpful in kick-starting academic investigation into this field. It was a critical time point for moving the field forward and it is getting to that tipping point where people are trying to move it forward at a lot faster pace.” (Kaplan)
- “I think that we have created something of a community. We have made some connections across different grantees and there has been some collaboration.” (Melichar)
- “To do something like this that is really spade work—and potentially the ground work for what really is something huge—took a lot of initiative and thoughtfulness on the Foundation’s part. I don’t see a lot of other people in that zone other than RWJF.” (Rubenstein)

AFTERWARD

Projects conducted through *Improving the Science of Continuous Quality Improvement Program and Evaluation* ended at the completion of the individual grants. Grantees continued with additional data analyses, prepared manuscript for publication, presented at conferences, and submitted proposals to other funders to expand the work begun under these grants.

Grantees are members of the developing ASQIRE community and, through that, may continue to develop connections among each other as well as with grantees of the national program *Evaluating Quality Improvement Training Programs* and other quality improvement researchers and practitioners.

Prepared by: Mary B. Geisz

Reviewed by: Kelsey Menehan and Molly McKaughan

Program Officer: Lori Melichar

Grant ID: CQI

Program area: Human Capital

APPENDIX 1

Recommendations from Mittman/Salem-Schatz Report

Recommendations to Address Challenges Associated With Designing and Managing CQI Program Evaluations

- Begin quality improvement initiatives with a clear articulation of RWJF’s goals and purposes.
- Use a systematic approach to program evaluation and planning.
- Encourage alternative paradigms to facilitate learning about quality improvement program implementation.
- Encourage a collaborative approach between program, evaluation, and site teams in line with program goals for learning.
- Make decisions about program tradeoffs explicit and put it in writing.³⁹

Recommendations to Address Data Collection and Measurement Challenges

- Consider whether additional data collection adds value when resources are stretched.
- Consider using local data collected by sites to assess the success of their CQI efforts for evaluation only in limited circumstances.
- Provide adequate financial and structural support for data collection and measurement activities.
- Promote the development and endorsement of nationally standardized measures.

Recommendations for Advancing the Field of CQI Research and Evaluation

- Enhance the legitimacy and use of a broad range of research and philosophy-of-knowledge paradigms.
- Encourage development and use of theoretical models, conceptual frameworks, and logic models to explain CQI initiatives.
- Support the development of methods, tools, and standards for the design, conduct, and reporting of CQI research and evaluations.

³⁹Those anticipating evaluation results should understand that results of evaluations may fall short of what they would ideally like to know. Where evaluation objectives are forfeited in order to improve program performance, funders and national program office staff members need to understand what that means in terms of information that can be shared with the field.

- Expand opportunities to publish and disseminate a broader range of CQI implementation and evaluation reports.
- Expand opportunities for health care professionals already engaged in CQI to be trained and encouraged to fully document their CQI processes and results in a manner that will support later analysis and synthesis.
- Support improved synthesis of findings from CQI research and evaluation efforts.

APPENDIX 2

List of Funded Projects

Project to Provide a Framework for the Identification, Classification, and Evaluation of Quality Improvement Initiatives

RAND Corporation (Santa Monica, Calif.)

Providing a framework for the identification, classification, and evaluation of research on continuous quality improvement

ID# 65113 (October 2008–December 2010) \$220,726

Providing a framework for the identification, classification, and evaluation of quality improvement initiatives

ID# 67890 (July 2010–January 2011) \$22,173

Project Director

Lisa V. Rubenstein, MD, MSPH

(310) 393-0411 x6303

lisar@rand.org

Projects to Develop New Quality Improvement Measures

Case Western Reserve University, Frances Payne Bolton School of Nursing (Cleveland)

Developing and conducting psychometric testing on a tool to measure systems thinking

ID# 65114 (October 2008–December 2010) \$99,938

Project Director

Mary A. Dolansky, PhD, RN

(216) 368-0568

mary.dolansky@case.edu

Case Western Reserve University School of Medicine (Cleveland)

Using a mixed methods approach to assess the settings in which interventions to improve quality are deployed in diverse health settings

ID# 65213 (October 2008–December 2010) \$84,711

Project Director

David G. Litaker, MD, PhD

(216) 791-3800 x4849

David.Litaker@case.edu

Cincinnati Children’s Hospital Medical Center (Cincinnati)

Developing a framework for optimizing contextual factors that impact quality improvement projects and their likelihood of success

ID# 65149 (October 2008–December 2010) \$173,881

Project Director

Heather C. Kaplan, MD

(513) 803-0478

heather.kaplan@cchmc.org

Johns Hopkins University Bloomberg School of Public Health (Baltimore)

Validating and testing feasible measures of health care continuous quality improvement team characteristics in relation to ICU patient outcomes

ID#65248 (October 2008–December 2010) \$155,376

Project Director

Jill A. Marsteller, PhD, MPP

(410) 502-0958

jmarstel@jhsph.edu

University of Medicine and Dentistry of New Jersey (New Brunswick, N.J.)

Enhancing the theory and practice of continuous quality improvement through a better understanding of organizational capacities

ID# 65242 (October 2008–December 2010) \$197,914

Project Director

Elizabeth C. Clark, MD, MPH

(732) 743-3239 x3

clarkeec@umdnj.edu

Projects Addressing Data Collection and Analysis Methodologies

Beth Israel Deaconess Medical Center (Boston)

Determining the reliability of a common tool used to measure continuous quality improvement

ID# 65121 (October 2008–September 2009) \$67,675

Project Director

Michael D. Howell, MD, MPH

(617) 632-7687

mhowell@bidmc.harvard.edu

Medical University of South Carolina (Charleston, S.C.)

Developing and testing tools for analyzing health care quality improvement data

ID# 65188 (October 2008–December 2010) \$52,728

Project Director

Jeffrey J. Borckardt, PhD

(843) 792-3295

borckard@musc.edu

Rutgers, the State University of New Jersey (Piscataway, N.J.)

Designing and implementing statistical methods for detecting changes in organizational performance

ID# 65181 (October 2008–December 2010) \$62,000

Project Director

Thomas O. Boucher, PhD, MBA, MS

(732) 445-3657

tboucher@rci.rutgers.edu

PROFILE LIST

Michael D. Howell, MD, MPH

Heather C. Kaplan, MD

Lisa V. Rubenstein, MD, MSPH