REVIEW ARTICLE

Evaluating the State of Quality-Improvement Science through Evidence Synthesis: Insights from the *Closing the Quality Gap* Series

Kathryn M McDonald, MM; Ellen M Schultz, MS; Christine Chang, MD, MPH

Abstract

Context: The *Closing the Quality Gap* series from the Agency for Healthcare Research and Quality summarizes evidence for eight high-priority health care topics: outcomes used in disability research, bundled payment programs, public reporting initiatives, health care disparities, palliative care, the patient-centered medical home, prevention of health care-associated infections, and medication adherence.

Objective: To distill evidence from this series and provide insight into the "state of the science" of quality improvement (QI).

Methods: We provided common guidance for topic development and qualitatively synthesized evidence from the series topic reports to identify cross-topic themes, challenges, and evidence gaps as related to QI practice and science.

Results: Among topics that examined effectiveness of QI interventions, we found improvement in some outcomes but not others. Implementation context and potential harms from QI activities were not widely evaluated or reported, although market factors appeared important for incentive-based QI strategies. Patient-focused and systems-focused strategies were generally more effective than clinician-focused strategies, although the latter approach improved clinician adherence to infection prevention strategies. Audit and feedback appeared better for targeting professionals and organizations, but not patients. Topic reviewers observed heterogeneity in outcomes used for QI evaluations, weaknesses in study design, and incomplete reporting.

Conclusions: Synthesizing evidence across topics provided insight into the state of the QI field for practitioners and researchers. To facilitate future evidence synthesis, consensus is needed around a smaller set of outcomes for use in QI evaluations and a framework and lexicon to describe QI interventions more broadly, in alignment with needs of decision makers responsible for improving quality.

Introduction

The quality of health care in the US is widely recognized as needing improvement. Indeed, as many as 50% of all patients, on average, may receive suboptimal care.[1-3] Yet quality is improvable, and efforts to make improvements are widespread.[1,4,5]

Just as medical science focuses on treating ailments and supporting the health of the human body through medical, surgical, pharmacologic, and preventive interventions, the science of quality improvement (QI) focuses on "treating" quality gaps and supporting optimal performance of the health care system through improvement interventions and quality monitoring. A key question for both medical and improvement science is how altering one part of a system—either the human body or the health care system—produces desired results. Additional questions relate to how interventions interact with the surrounding environment and circumstances (the context of change) and how delivery of the intervention (implementation of change) has an impact on effectiveness. Many of the tools of medical research that were tailored to answer such questions have also been applied to improvement science, including systematic reviews and meta-analyses.

In 2004, the Agency for Healthcare Research and Quality (AHRQ) launched a collection of systematic reviews on QI strategies related to high-priority chronic conditions (eg, diabetes, asthma, hypertension), practice areas (eg, prevention of health care-associated infections, antibiotic prescribing behavior), and processes (eg, care coordination) identified by the Institute of Medicine.[6-12] AHRQ followed this collection with a new series of eight evidence reports—*Closing the Quality Gap: Revisiting the State of the Science*—to continue the focus on improving the quality of health care, including current efforts to reward high-quality care through measurement and reporting as well as key tenets of health care reform legislation passed under the Patient Protection and Affordable Care Act.[13] In addition, through two cross-topic synthesis projects,[14,15] the new series of reports also sought to illuminate broader lessons about the state of QI science by aggregating evidence in a qualitative way across the sample of topics included in the series.

This article builds on that synthesis, summarizing the "state of the science" for the effectiveness, implementation decision factors, and evidence base of the QI field on the basis of findings from the most recent *Closing the Quality Gap* series of topic reports.

Methods
Series Topics

The *Closing the Quality Gap* series included eight topics selected by leaders in AHRQ for their relevance to high-priority populations, settings, and processes,[4] and to provisions of the Affordable Care Act (Table 1). Selected topics were also ripe for systematic review and expected to yield

Kathryn M McDonald, MM, is a Senior Scholar and Executive Director of the Center for Health Policy/Center for Primary Care and Outcomes Research at Stanford University in CA. E-mail: kathryn.mcdonald@stanford.edu. **Ellen M Schultz, MS,** is a Project Coordinator at the Center for Health Policy/Center for Primary Care and Outcomes Research at Stanford University in CA. E-mail: emschultz@stanford.edu. **Christine Chang, MD, MPH,** is a Medical Officer at the Center for Outcomes and Evidence of the Agency for Healthcare Research and Quality in Rockville, MD. E-mail: christine.chang@ahrq.hhs.gov.

REVIEW ARTICLE

Evaluating the State of Quality-Improvement Science through Evidence Synthesis: Insights from the Closing the Quality Gap Series

actionable evidence for patients, practitioners, health systems, and policy makers.

We mapped these topics to three core approaches ("3 Is") for achieving improvements, as noted by health care systems researcher Victor Fuchs,[16] who said that real reform "requires changes in the organization and delivery of care that provide physicians with the *information, infrastructure,* and *incentives* they need to improve quality and control costs." In today's complex health care system, these leverage points for improvement apply beyond the physician to include other clinicians, systems managers, and patients themselves. The set of topics selected for the series address each of these three core approaches (see Table 1).

Topic Reviews

Each topic was reviewed by a team from an AHRQ Evidence-based Practice Center (EPC) using a standard methods guide.[17] Complete details of review methods for each topic are available in the individual topic reports.[18-25] A brief summary is presented here. In conjunction with topic-specific technical expert panels, team members of each EPC developed a set of key questions to guide their review. The EPC teams searched a wide variety of literature databases, including at a minimum MEDLINE, and an average of 25 years of literature for each topic (range, 5 to 65 years). They identified relevant articles through multiple rounds of review and abstracted detailed information from each included study. All studies were evaluated for quality and potential bias using a standard protocol. Likewise, when reported and applicable, evaluations of strength of evidence across studies also followed standard methods.[17]

Cross-Topic Synthesis

Results presented in this article are based on the eight series topic review reports.[18-25] We initially provided common guidance to each topic review team for the series to facilitate cross-topic synthesis. Then we reviewed the evidence presented in the reports, including tables and text, to identify cross-cutting themes, take-home lessons, common challenges, and evidence gaps as they relate to the science of QI. Thus, this synthesis is based on comparisons across the series topic reports rather than on primary studies reviewed in those reports. We did not perform quantitative meta-analyses, but instead focused on qualitative synthesis to provide insight into the field of QI. Additional discussion of topic-specific findings and implications for key stakeholder audiences may be found in the series summary report[15] and an accompanying methods report.[14]

Key Questions

We developed a set of series key questions to guide evidence synthesis across series reports. These key questions focus on the "state of the science" for three core aspects of QI: effectiveness, implementation decision factors, and evidence. The key question areas are as follows:

1. What is the state of the evidence for the *effectiveness* of QI activities? What outcomes have been examined in evaluating effectiveness? What is known about the benefits and harms of particular types of QI strategies or targets?
2. What is the state of the science for factors of likely importance to those individuals and organizations deciding whether and how to *implement* QI interventions? What is known about the role of context and implementation approaches/challenges in QI activities? What is known about the impact of QI activities on disparities or vulnerable populations?
3. What is the state of QI and implementation science *evidence?* What gaps exist in the quality of evidence or in methods for evidence synthesis?

We summarized evidence of effectiveness—both benefits and potential harms—for the series topics (excluding disability outcomes, which focused on use of outcomes and did not address effectiveness) and considered the role of outcomes choice in effectiveness evaluations. We also examined evidence of effectiveness for QI strategies by type, using a taxonomy of improvement strategies developed for the first *Closing the Quality Gap* series.[10] We grouped these strategies by the intervention target—patients, clinicians, or systems/organizations—to further analyze evidence of effectiveness.

Table 1. *Closing the Quality Gap* series topics and primary focus for leverage to improve quality of care	
Topic	**Focus**
Information: Providing information about outcomes used in evaluating health care quality	
Disability outcomes[18]	Identify outcomes measures used in quality-focused research involving people with disabilities
Incentives: Influencing improvement through payment changes and quality monitoring	
Bundled payment[20]	Examine the influence on organizations of adopting payment bundling as an approach for paying for care (contrasted with fee-for-service models), and how organizational response to such new incentives either enhances or deters health care quality, including efficiency
Public reporting[23]	Understand how public reporting of health care quality information affects behaviors of people and organizations in ways that potentially improve the quality of care received by patients
Infrastructure: Changing delivery infrastructure to improve quality of care	
Disparities[22]	Examine the benefits and harms of quality-improvement interventions aimed specifically at reducing disparities in care
Palliative care[19]	Examine the impact on health care quality of various aspects of palliative care, including palliative care delivered in hospice and nursing homes
Patient-centered medical home (PCMH)[25]	Understand whether and how implementation of a comprehensive PCMH improves care overall for the full population of patients served by a health care organization
Health care-associated infections (HAI)[21]	Examine effectiveness of quality-improvement efforts aimed at improving adherence with evidence-based HAI-prevention strategies, including at ambulatory surgical centers, dialysis centers, and long-term care facilities
Medication adherence[24]	Address both the efficacy and effectiveness of interventions designed to improve medication adherence for adults with chronic conditions, including system and policy-level interventions

REVIEW ARTICLE

Evaluating the State of Quality-Improvement Science through Evidence Synthesis: Insights from the Closing the Quality Gap Series

Table 2. Summary of key findings for intervention-focused series topics[a]			
Topic	Evidence of improvement	Mixed or no evidence of improvement	Potential harms or unintended consequences
Bundled payment	Small decreases in health care spending (≤10%) and health care utilization (5%-15% reduction)	Mixed impact on quality measures (variable magnitude and direction of effects)	Single-setting bundled payment programs resulted in care shifting to other settings Few studies addressed other potential harms
Public reporting	Hospital-level reporting shows decreased mortality Health plan and long-term-care-level reporting shows improved pain, pressure ulcers, patient/family satisfaction Clinicians and health care organizations responded to public reporting by offering new services, changing policies, and by increasing quality-improvement activities	Few patients use public reports to make health care decisions; reports lack relevance or clarity or are unavailable when needed	Overall, evidence of no harm outweighed evidence of harm Mixed results and low-quality evidence about impact of public reporting on patient access Some evidence that public reporting in long-term care led to changes in coding practices and readmitting patients before reporting assessment Evidence refuted claims that public reporting leads surgeons or health care organizations to withdraw from the market or that public reporting is associated with declines in quality of unmeasured aspects of care (crowding out)
Disparities	A single study showed reduced disparity in HbA$_{1c}$ testing among black vs white patients with a disease management and patient education program Limited evidence of amplified effects of collaborative care and language-concordant patient education strategies in vulnerable populations	Insufficient evidence for changes in disparity after quality-improvement interventions; few studies addressed the research question	No studies addressed potential harms
Palliative care	Interventions targeting: Pain: improved pain-related outcomes Coordination: improved patient/family satisfaction Communication/decision making: improved health care utilization	Interventions targeting: Pain: no improvement in QOL or health care utilization Coordination: no improvement in QOL, symptoms, or health care utilization Communication: no improvement in patient/family satisfaction or health care utilization No interventions using only clinician-focused strategies were effective	Not examined
Patient-centered medical home (PCMH)	Small improvements in patient and staff experiences (satisfaction with care, perception of coordination) Decreased use of Emergency Department by older adults Small positive effects on delivery of preventive services	No decrease in hospital admissions for older adults No evidence of cost-savings with PCMH	Limited evidence from two studies found that when program costs were considered, the overall cost of care was greater for the PCMH intervention No other evidence on potential harms from PCMH reported in included studies
Health care-associated infection (HAI)	Some combinations of strategies show improved adherence to best practices and lower infection rates	Organizational change and clinician education alone did not improve adherence or infection rates Insufficient evidence to draw conclusions about improvements in cost savings or return on investment	No studies addressed potential harms
Medication adherence	Adherence improved with policy-level change decreasing patients' out-of-pocket costs and several other patient-focused strategies	Decreasing patients' costs did not improve adherence with inhaled corticosteroids Only a subset of studies showing improved adherence also improved other disease-specific clinical outcomes Studies of medication adherence interventions rarely examined impacts on health care utilization or costs; evidence is inconclusive	Few studies addressed potential harms

[a] The disability outcomes report did not evaluate the effectiveness of quality-improvement interventions and therefore is not included in this table.
HbA$_{1c}$ = glycosylated hemoglobin; QOL = quality of life.

REVIEW ARTICLE

Evaluating the State of Quality-Improvement Science through Evidence Synthesis: Insights from the Closing the Quality Gap Series

To examine the state of the science regarding factors likely to inform implementation decisions, we summarized findings from each report that relate to the context of QI implementation or evaluation, implementation approaches and challenges, and the impact of QI efforts on health disparities or on vulnerable populations. Finally, we evaluated the state of the science on the basis of the entire evidence base, summarizing common challenges encountered by the EPCs and identifying gaps in the evidence and in systematic review methods applied to improvement and implementation science.

Results

Key Question 1: Effectiveness of Quality-Improvement Strategies

Table 2 summarizes key findings about the effectiveness of QI efforts for each of the seven series topics that evaluated interventions. Authors of all seven topics found mixed results, with evidence of benefit for some outcomes but not for others. For example, the bundled payment review found evidence that the impact of payment bundling on quality of care depended on the quality measure evaluated (Table 2). The medication adherence review authors found variability in how adherence was defined, and they noted that only a subset of studies reporting improved adherence also showed improvements in other outcomes.

Six reports sought information about potential harms associated with QI interventions (Table 2). Potential harms were evaluated most often for the incentive-based interventions (bundled payment, public reporting), whereas harms were rarely addressed in the literature reviewed for the infrastructure-focused intervention topics (disparities, patient-centered medical home, health care-associated infections, medication adherence). Although the potential for harm from public reporting was widely discussed, the review authors found only limited evidence examining whether harm actually occurred and concluded that evidence of no harm outweighed evidence of harm. The bundled payment review found consistent evidence that single-setting bundled payment programs resulted in care shifting to other settings, but few other potential harms were examined. The review authors noted that most current bundled payment programs are now administered across settings, which is expected to reduce incentives for care shifting.

The disability outcomes review identified 71 different outcomes measures used in evaluating health care for disabled populations. Many of these assessed similar concepts, including health, quality of life, functioning, and patient experience, but used different definitions, tools, and measurement scales. The review authors also noted that researchers' perspective—whether trained and practicing in medicine, rehabilitation, or social services—had a profound impact on the ways in which care and life goals were conceptualized for people with disabilities, influencing their choice of outcomes for evaluation.[18]

Across the series topics, most QI interventions were multifaceted, using more than one type of improvement strategy (Table 3). There was greater evidence of effectiveness for systems-focused strategies than for either clinician- or patient-focused strategies. However, most evidence of systems-focused strategies related to organizational change, which can encompass many different kinds of activities.[10] For most topics examined, clinician-focused strategies were generally less effective than patient-focused strategies, with the exception of interventions aimed at improving clinician adherence to strategies to prevent health care-associated infections. Among the patient-focused strategies, patient education often showed benefit.

In contrast, evidence of effectiveness was mixed for patient and clinician reminder systems and for audit and feedback strategies (Table 3). The latter strategies can be patient-focused when aimed at influencing consumers' decisions about where to seek care, such as through public reporting of quality information. These strategies can be clinician-focused

Table 3. Effectiveness of quality-improvement interventions by strategy type and target[a]							
Strategy type	Bundled payment	Public reporting	Disparities	Palliative care	PCMH	HAI	Medication adherence
Target: Patients							
Education			■	■		□	■
Promotion of self-management			□	■		□	■
Reminder system						□	■
Audit and feedback		□					
Target: Clinicians							
Education			□	□		■	
Reminder system				□		■	
Facilitated relay of clinical data				□		□	□
Audit and feedback		■		□		■	
Target: Systems							
Organizational change			□	■	■	■	■
Financial, incentives, regulation, and policy	■						
Audit and feedback		■		■		□	■

[a] The disability outcomes report did not evaluate the effectiveness of quality improvement interventions and therefore was not included in this table.

HAI = health care-associated infection; PCMH = patient-centered medical home; ■ = intervention type has been examined, and there is evidence of effectiveness; □ = intervention type has been examined, but there is no evidence of effectiveness.

REVIEW ARTICLE

Evaluating the State of Quality-Improvement Science through Evidence Synthesis: Insights from the Closing the Quality Gap Series

when aimed at motivating clinicians to make changes in their practice on the basis of their performance on quality measures. Alternatively, the strategies can be system-focused if intended to influence organizations' practices or motivate QI efforts. Four reports found evidence related to audit and feedback strategies, showing that they were not effective when targeting patients but were generally effective when targeting clinicians and organizations.

Differences in outcomes seen across topics may reflect topic-specific differences in the locus of control, contextual factors, variable adaptation of intervention components, interaction between intervention components, and underlying barriers to improved performance. Reviews typically found limited details about the presumed mechanism of an intervention for influencing behavior (sometimes referred to as the logic model), limiting synthesis-based insights about which interventions are effective and why.

Key Question 2: Quality-Improvement Implementation Decision Factors

Many of the reports examined three key drivers of QI implementation decisions: the role of context, implementation approaches and challenges, and the impacts of QI efforts on vulnerable populations or health care disparities (Table 4). In assessing contextual factors to determine reasons for amplification or dampening of the effect of an intervention, both the bundled payment and public reporting reviews found evidence that these incentive-based strategies were more effective when financial pressures were greater, such as in competitive markets (public reporting), and in for-profit or financially stressed hospitals (bundled payment). Other reports of contextual factors varied greatly in the type of factors examined and their use in the primary studies, ranging from economic

considerations to patient characteristics (disease severity, age, insurance coverage, health needs) and organizational characteristics (leadership, change, resource availability). All five series reports that examined the role of context in some manner (Table 4) found that information on contextual factors was often lacking, incompletely described, or noted only anecdotally.

Aside from a specific focus on implementation in two reports (health care-associated infections, bundled payment), and explicit exclusion of implementation studies in the public reporting review (because of lack of outcomes available in relevant studies), the remaining five reviews had limited coverage of implementation approaches and challenges (Table 4). Two reports noted challenges related to clinician resistance to interventions (bundled payment, palliative care) and two reports identified resource issues (disparities, patient-centered medical home). Several reports (patient-centered medical home, health care-associated infections, medication adherence) sought information about approaches used to enable implementation (eg, toolkits, collaborative learning).

Four reports examined the impact of QI efforts or choice of evaluation outcomes on health disparities or vulnerable populations (Table 4). Although the available literature was limited, the disparities report found some promise for reducing disparities in health outcomes among racial minorities using collaborative care and targeted patient education interventions. Racial and ethnic minorities were the most widely studied vulnerable populations across the topics.

Key Question 3: State of Quality-Improvement Evidence

The EPC teams conducting the topic reviews encountered several common challenges that limited their ability to synthesize evidence across studies and to address their research questions. Many of these challenges stemmed from limitations in the primary studies. Members of the EPCs for all eight topics observed great heterogeneity in choice and definition of outcomes used for QI evaluations. They also noted study design weaknesses and incomplete reporting of key details such as intervention

design and its theoretical basis, contextual factors and impact on outcomes, intervention components, and comparators.

Across the series, just a handful of conclusions were based on moderate or high strength of evidence (the confidence that a conclusion reflects a true effect). They were as follows: reducing the patient's out-of-pocket costs improved medication adherence (moderate strength of evidence), hospital-level public reporting decreased mortality rates (moderate strength of evidence), and public reporting stimulated improvement in competitive markets and among low performers (high strength of evidence). The strength of evidence for most other research questions addressed across the series topics was low or inconclusive.

These limitations in the primary studies created challenges in adapting systematic review methods to the QI literature. The heterogeneity in outcomes, coupled with the complexity of multifaceted, systems-level interventions typical of the QI literature, limited the ability of the EPC teams to quantitatively synthesize results across studies. They instead summarized evidence qualitatively, grouping evidence by particular disease groups, settings, outcomes, or intervention components. Ambiguity around use of key terms in the primary studies (eg, QI itself, as well as some topic-specific terms such as *medical home* and *palliative care*) complicated development of search strategies. Other systematic review challenges included assessment of the body of evidence across heterogeneous studies and the lack of statistical or other approaches to synthesize across a diversity of study designs, intervention components, implementation factors, contextual factors, and outcomes.

Some challenges encountered may positively reflect characteristics of QI evidence. Whereas heterogeneity in QI strategies presented difficulties in synthesis and drawing conclusions, this also reflects the variety of strategies used in practice that are likely to be relevant to decision makers. Similarly, heterogeneity in outcomes offers many different lenses through which to view quality of care. Furthermore, despite challenges, the methodologic quality of the evidence base has improved, as noted by the authors of the report on health care-associated infections.[21]

> … conclusions were based on moderate or high strength of evidence were: reducing the patient's out-of-pocket costs improved medication adherence, hospital-level public reporting decreased mortality rates, and public reporting stimulated improvement in competitive markets and among low performers …

Evaluating the State of Quality-Improvement Science through Evidence Synthesis: Insights from the Closing the Quality Gap Series

All reports found a body of evidence to synthesize. Most reports included various study types to complement evidence from controlled trials, providing additional detail that improved the usefulness of the reports.

Discussion

This *Closing the Quality Gap* series systematically reviewed and synthesized evidence relating to eight QI topics. Although far from inclusive of all QI efforts, the eight topics included within this series represent a sample of the range of topics, populations, settings, strategies, and improvement targets within the broader universe of QI science; they cover three critical leverage points for improving care: information, incentives, and infrastructure.[16]

Table 4. Summary of findings by topic on implementation decision factors			
Topic	Context	Implementation approaches and challenges	Impacts on vulnerable populations and disparities
Disability outcomes	Not applicable	Reviewers emphasize that choice of outcomes can lead to problems if poorly matched to population needs and values	Disabled populations rarely included in studies with nondisabled patients
Bundled payment	Some weak evidence that bundled payment decreased health care utilization more among for-profit providers compared with not-for-profit providers and at hospitals under greater financial pressure	Some survey evidence that new bundled payment systems faced initial resistance from clinicians For each bundled payment study or set of studies, topic review includes a section on reported implementation challenges	Not examined
Public reporting	Strong evidence that public reporting leads to improvements in competitive markets and among low performers	Not examined	One study found increased disparity between white and black or Hispanic patients in the receipt of coronary artery bypass graft surgery with public reporting
Disparities	Not examined	Reviewers noted that some interventions required substantial programmatic and implementation resources	Few studies have examined QI strategies as a way to reduce health disparities Limited evidence suggests some reduction in health outcome disparities with collaborative care and targeted patient education interventions, particularly among racial minorities
Palliative care	Some evidence supported the effectiveness of both integrative and consultative models for delivering palliative care	Some studies reported challenges with clinician uptake of interventions, as well as difficulties with recruitment or retention of participants in QI activities	Not examined
Patient-centered medical home (PCMH)	Payment models used to support PCMH implementation varied widely, including receipt of external study funding, capitation payments, enhanced fee-for-service, and a hybrid approach. Fewer than half of studies described their payment model	Horizon scan identified a number of planned formative evaluations to identify factors associated with successful implementation. Cost to practice noted as a factor for study Implementation usually included formal learning collaboratives or collaborative program planning for practice team for members to learn about the new intervention (19 of 22 studies), and audit and feedback strategies were often tied to QI (13 of 22 studies)	Not examined
Health care-associated infection (HAI)	Wide variety of contextual factors reported. Three most commonly reported factors were availability of implementation materials, unit-level changes in responsibilities, and unit-level leadership	QI strategy was defined as the implementation strategy (eg, clinician education regarding an HAI preventive intervention). Implementation materials were described as a contextual feature for some interventions	Not examined
Medication adherence	Not examined	Minimal or unclear information available on implementation (eg, organizational learning strategies, use of implementation toolkits, or fidelity to intervention protocol)	Interventions generally had a positive impact on medication adherence for most vulnerable populations examined. These populations were typically defined by race-ethnicity and medical condition

QI = quality improvement.

REVIEW ARTICLE

Evaluating the State of Quality-Improvement Science through Evidence Synthesis: Insights from the Closing the Quality Gap Series

Individually, each of the series reviews offers detailed information that can help inform QI efforts and decisions related to its respective topic. Viewing the evidence together across series reports revealed broader insights. For example, the finding that both the incentive-based improvement topics (bundled payment and public reporting) were sensitive to the market context—competitiveness of the health care market, financial pressure on delivery organizations—suggests that particular attention should be paid to the market and financial context of any incentive-based improvement efforts. Context is likely also important to consider for information and infrastructure-based improvement efforts. The disability outcomes reviewers observed that the professional background of researchers influenced their conception of how to evaluate interventions for disabled populations, highlighting the relevance of the evaluation context, especially choice of outcomes.

Looking across topics, the series also found evidence supporting the effectiveness of broader types of intervention strategies, in particular organizational change. Although specific studies varied with respect to the kinds of organizational change implemented (eg, collaborative care, patient-centered medical home, case management) and ways in which organizational change was combined with other intervention strategies, these results suggest that this is likely an important component of many effective QI interventions.

Additional patterns of effectiveness became apparent when examining the target for improvement strategies. Public reporting, an example of an audit and feedback strategy, was generally effective in changing clinician and organizational behaviors, but not patients' behavior. Qualitative evidence included in the public reporting review supported this finding. Interventions that focused solely on clinicians as a target group tended to demonstrate less benefit, with the exception of the topic of health care-associated infections.

The teams from the EPCs also identified a gap in examination and reporting of potential harm from QI activities. Although examination of side effects of medical therapies is expected in the medical literature, the reviews revealed that few studies of QI efforts have addressed the potential for unintended negative consequences. Among the series topics, public reporting had received the most attention toward potential harms, but even for this topic, the reviewers found that the potential for harm was discussed far more often than it was evaluated. This gap in QI evidence is ripe for development, and it may require guidelines for evaluating and reporting harms that may be far-reaching or that may occur well after the initial intervention.

In addition to these insights, synthesis of evidence across series topics also sheds light on the "state of the science" for the QI field itself. The common challenges experienced by the teams from the EPCs highlight areas where additional methods

---

Table 5. Example typology for quality-improvement interventions[a]

Intervention target: The target refers to the person, people, health system, or policy to which intervention activities are directed. Interventions may directly target providers, patients, aspects of a health system, health policies, or some combination of these four.[24]

Intervention agent: An intervention agent is the person, people, or technology used to deliver the intervention. Examples of possible intervention agents include physicians, nurses, pharmacists, case managers, multidisciplinary teams, or family members. Some interventions may have more than one agent delivering an intervention or a part of an intervention.[24]

Mode of delivery: The mode of delivery refers to the manner by which the agent delivers the intervention. For example, interventions may be delivered face-to-face, by telephone, with print materials, or by computer, DVD, video, or CD/audio. Like intervention target and agent, an intervention may have more than one mode of delivery.[24]

Intensity of intervention: Intensity refers to the total amount of time an intervention lasts, taking into account the duration and number of all individual sessions (eg, five 30-minute sessions or one 60-minute session).[24]

Duration of intervention: In contrast to intensity, the duration of an intervention is a description of the total length of calendar time over which any series of individual sessions are delivered. Two interventions may have the same total intensity (eg, five 30-minute sessions) but be spread out over different total durations of time (eg, one over 1 month, another over 1 year).[24]

Intervention components: Frequently, multiple components are used to create a multifaceted intervention strategy.[24] One taxonomy developed for the original *Closing the Quality Gap* series specifies nine types of improvement strategies:[10]

    Clinician reminder systems
    Facilitated relay of clinical data to clinicians
    Audit and feedback
    Clinician education
    Patient education
    Promotion of self-management
    Patient reminders
    Organizational change
    Financial, regulatory, or legislative incentives

Implementation context: The circumstances under which the QI intervention is implemented. One set of contextual factors adapted from the patient safety field[26] lends structure and common language possibilities:

Structural organizational characteristics: organization size, location, financial status, academic status, complexity, volume, existing quality infrastructure, space/physical environment, use of information technology, physician ownership, and the dates of study.

External factors: regulatory environment, payments and penalties, local sentinel event, marketplace competition, competing demands.

Culture, teamwork, leadership: each of these three factors can be examined at the organizational level or unit level.

Implementation and management tools: includes use of specific improvement strategies targeted at clinicians or staff (such as education, audit and feedback, or financial incentives) staff education, designated staff time to implement change, designating an internal or external person responsible for implementation, local tailoring or use of an iterative process, help desk support, extent of project management, implementation timeline, and implementation process (one unit at a time or all at once).

[a] Adapted from the medication adherence review,[24] the original Closing the Quality Gap series,[10] and a patient safety context report from the Agency for Healthcare Research and Quality.[26]

CD = compact disk; DVD = digital video disk; QI = quality improvement.

REVIEW ARTICLE

Evaluating the State of Quality-Improvement Science through Evidence Synthesis: Insights from the Closing the Quality Gap Series

or conceptual development is needed. Inconsistencies in how interventions are described in the literature point to the need for an underlying framework and lexicon to describe QI interventions. Although a framework and terminology must be flexible enough to cover the diverse universe of QI strategies, consistent use of a common set of terms would help facilitate synthesis of results across studies, as was done in the cross-topic review presented in this article. Table 5 presents an example of a typology used to describe improvement interventions, adapted from the medication adherence review.[24] As Table 3 demonstrated, combining one element of this typology—the intervention target—with the taxonomy of improvement strategies used in the original *Closing the Quality Gap* series[10] provided insights that apply across topics and that were not readily apparent without this structure. Reaching consensus around a common framework and lexicon for QI science requires further development, but the approach demonstrated in this synthesis presents a useful starting place in that endeavor.

The evidence base is growing regarding the importance of context for quality and patient safety topics,[26-28] yet all five series reports that examined the role of context found that implementation context was rarely described in the QI literature. The teams from the EPCs recommended that contextual factors be more frequently and robustly measured and reported. To accomplish this will require development of reliable and valid measures of such factors, but at this early stage of exploration, little is known about which contextual factors are important to measure, and how to do so. Thus, filling this knowledge gap will require iterative measure development, measurement, research, and refinement of the measures. Each of these steps will contribute valuable knowledge to the field. Table 5 includes a starter set of contextual factors, adapted from the patient safety field,[26] that can help lend structure and a common language to future work around implementation context. These context factors also map well to the Consolidated Framework for Implementation Research.[29,30]

This "meta" review evinces the promise of scaling up knowledge across topics

through a structured qualitative synthesis, in this case relying on a common conceptualization of different levers (information, infrastructure, and incentives) for influencing behavior change to improve clinical and economic outcomes, a typology of QI strategies and contexts (Table 5), and attention to potential harms and vulnerable populations. To foster useful description and synthesis, we also recommend extending a framework acronym that is commonly applied to systematic reviews of clinical interventions to the needs of QI evaluation. Thus, PICOTS (population, intervention, comparator, outcomes, timing, setting) becomes PLICCOTS, adding "L" for logic model, and "C" for context. The overarching question for QI studies is then: for a defined population, what is the logic argument for a complex intervention working better than its comparator in a given context to produce outcomes (of interest to QI) within a time period and setting?

All eight reviews in this series were limited in their ability to synthesize the available evidence and draw conclusions across studies in part because of the extreme heterogeneity in the outcomes reported and ways in which those outcomes were measured, which highlights the need for more consistent outcomes measurement. Developing a set of consensus-based, clearly defined, and fully specified outcomes measures for use in QI research would help facilitate evidence synthesis by enhancing comparability across studies, although use of standardized measures must be balanced with the need to tailor choice of outcomes to the goals of particular QI efforts or research studies. Some efforts to develop core measure sets and harmonize quality measures are underway[31-35] and hold promise for advancing the state of QI science if the resulting consensus-based measures are widely used. For example, the US Department of Health and Human Services (HHS), as a part of the National Strategy for Quality Improvement in Health Care, is developing measure selection processes with public input and the goal of aligning measures across new and existing HHS programs and focusing on patient outcomes and patient experience of care. Because of this effort, multiple measures for blood pressure control in use across

HHS were identified, and in the future a consensus-developed set of measures will be used across all HHS programs.[34]

Evidence synthesis by the series EPCs was further hampered by limitations in applying study designs and systematic review methods that were developed for evaluating clinical interventions to the kinds of multifaceted, context-dependent, systems-level interventions and implementation approaches typical of the QI field. However, the topic teams did explore various approaches to improve the relevance of the reports for decision makers. For example, the public reporting review included qualitative research to complement the quantitative studies, and the patient-centered medical home review included a horizon scan to inform decision makers about ongoing research. Because of the context-dependent nature of QI interventions, other complementary methods may inform questions related to policy and practice, and may provide information for better decision making. These methods could potentially help address the diversity of intervention components, implementation factors, and context. Advances could include qualitative research synthesis techniques, exploration of methods to systematically identify and assess gray literature, and exploration of methods to assess and incorporate a variety of study designs. Further methodologic attention to meta-analytic approaches is also needed to achieve sufficient statistical power with relatively few intervention units (eg, hospitals, clinics, health systems) for organization-level interventions. Although it is beyond the scope here to describe specific methods, the choice of method will depend on the anticipated use of the review, the type of questions asked, underlying assumptions, and breadth and depth of the proposed review. Overall, the preponderance of low strength of evidence findings and limited information on additional considerations of interest to local decision makers (eg, context, implementation approaches/challenges, vulnerable population impact)

> **The evidence base is growing regarding the importance of context for quality and patient safety topics, yet all five series reports that examined the role of context found that implementation context was rarely described in the QI literature.**

REVIEW ARTICLE

Evaluating the State of Quality-Improvement Science through Evidence Synthesis: Insights from the Closing the Quality Gap Series

found across the eight series reports speaks to the immaturity of the QI and implementation science fields. In these fields, decision-salient research questions and standards for robust and complementary study design continue to evolve.

Although synthesizing evidence across the series topics provided valuable insight into the state of QI science, the eight topics in the series represent just a sample of the QI field. Findings from this synthesis can help guide future QI efforts and suggest directions for future research but do not represent conclusive evidence of effectiveness or associations between particular strategies and other important factors. In addition, findings reported in this synthesis are presented in broad terms; much detail about the particular populations, settings, outcomes, and strategies included in the primary studies is omitted for the sake of highlighting conclusions that are applicable across major portions of the health care system. The individual topic reports provide much greater granularity in their findings and should be consulted to interpret particular topic-specific findings.

Conclusion

This series synthesis highlights the value in expanding our view from the level of individual improvement efforts to examine effectiveness of QI strategies across initiatives, topics, and targets. Limitations in the literature encountered by the EPCs point to areas in need of more rigorous standards for study design and reporting, methodologic weaknesses in need of further development, and research questions ripe for exploration. The findings also highlight common challenges limiting much of the QI literature, in particular, the lack of consensus around key outcomes important for evaluating QI effectiveness, gaps in analyzing other factors important to decisions about implementing a particular QI strategy, and weaknesses in study design and analytic methods. Using these challenges and methodologic weaknesses to generate practical and scientifically sound solutions can help guide future research efforts and the development of the QI field. ❖

References

1. Institute of Medicine. Best care at lower cost: the path to continuously learning health care in America. Washington, DC: The National Academy of Sciences; 2012.
2. Mangione-Smith R, DeCristofaro AH, Setodji CM, et al. The quality of ambulatory care delivered to children in the United States. N Engl J Med 2007 Oct 11;357(15):1515-23. DOI: http://dx.doi.org/10.1056/NEJMsa064637
3. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003 Jun 26;348(26):2635-45. DOI: http://dx.doi.org/10.1056/NEJMsa022615
4. Adams K, Corrigan JM, editors. Priority areas for national action: transforming health care quality. Washington, DC: The National Academies Press; 2003.
5. Committee on Quality of Health Care in America, Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington, DC: The National Academies Press; 2001 Mar 1.
6. Bravata D, Sundaram V, Lewis R, et al. Closing the quality gap: a critical analysis of quality improvement strategies (vol 5: asthma care). Technical reviews no. 9.5. Rockville, MD: Agency for Healthcare Research and Quality; 2007 Jan. AHRQ Publication No. 04-005101-5.
7. McDonald K, Sundaram V, Bravada D, et al. Closing the quality gap: a critical analysis of quality improvement strategies (vol 7: care coordination). Technical reviews no. 9.7. Rockville, MD: Agency for Healthcare Research and Quality; 2007 Jun. AHRQ Publication No. 04(07)-0051-7.
8. Ranji SR, Shetty K, Posley KA, et al. Closing the quality gap: a critical analysis of quality improvement strategies (vol 6: prevention of healthcare-associated infections). Technical reviews no. 9.6. Rockville, MD: Agency for Healthcare Research and Quality; 2007 Jan. AHRQ Publication No. 04(07)-0051-6.
9. Ranji SR, Steinman MA, Shojania KG, et al. Closing the quality gap: a critical analysis of quality improvement strategies (vol 4: antibi- otic prescribing behavior). Technical reviews no. 9.4. Rockville, MD: Agency for Healthcare Research and Quality; 2006 Jan. AHRQ Publication No. 04(06)-0051-4.
10. Shojania KG, McDonald KM, Wachter RM, Owens DK. Closing the quality gap: a critical analysis of quality improvement strategies (vol 1: series overview and methodology). Technical reviews no. 9.1. Rockville, MD: Agency for Healthcare Research and Quality; 2004 Aug. AHRQ Publication No. 04-005101.
11. Walsh J, McDonald KM, Shojania KG, et al. Closing the quality gap: a critical analysis of quality improvement strategies (vol 3: hypertension care). Technical reviews no. 9.3. Rockville, MD: Agency for Healthcare Research and Quality; 2005 Jan. AHRQ Publication No. 04-0051-3.
12. Shojania KG, Ranji SR, Shaw LK, et al. Closing the quality gap: a critical analysis of quality improvement strategies (vol 2: diabetes care). Technical reviews no. 9.2. Rockville, MD: Agency for Healthcare Research and Quality; 2004 Sep. AHRQ Publication No. 04-0051-2.
13. Timeline for health care reform implementation: system and delivery reform provisions [monograph on the Internet]. New York, NY: The Commonwealth Fund; 2010 Apr 1 [cited 2012 Aug 21]. Available from: http://mobile.commonwealthfund.org/~/media/Files/Publications/Other/2010/Timeline%20System%20Reform_040110_v5_rev%2051310.pdf.
14. McDonald KM, Chang C, Schultz E. Through the quality kaleidoscope: reflections on the science and practice of improving health care quality. Closing the quality gap: revisiting the state of the science. Methods research reports. Rockville, MD: Agency for Healthcare Research and Quality; 2013 Feb. AHRQ Publication No. 13-EHC041-EF.
15. McDonald KM, Chang C, Schultz E. Summary report. Closing the quality gap: revisiting the state of the science. Rockville, MD: Agency for Healthcare Research and Quality; 2013 Jan. AHRQ Publication No. 12(13)-E017.
16. Fuchs VR. The proposed government health insurance company—no substitute for real reform. N Engl J Med 2009 May 28;360(22):2273-5. DOI: http://dx.doi.org/10.1056/NEJMp0903655
17. Methods guide for effectiveness and comparative effectiveness reviews [Web page on the Internet]. Rockville, MD: Agency for Healthcare Research and Quality; 2012 Apr 9 [cited 2012 Jul 26]. Available from: http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318.
18. Butler M, Kane RL, Larson S, Jeffery MM, Grove M. Closing the quality gap: revisiting the state of the science (vol 7: quality improvement measurement of outcomes for people with disabilities). Evidence reports/technology assessments, no. 208.7. Rockville, MD: Agency for Healthcare Research and Quality; 2012 Oct. AHRQ Publication No. 12-E013-EF.
19. Dy S, Aslakson R, Wilson R, et al. Closing the quality gap: revisiting the state of the science (vol 8: improving health care and palliative care for advanced and serious illness). Evidence reports/technology assessments, no. 208.8. Rockville, MD: Agency for Healthcare Research and Quality; 2012 Oct. AHRQ Publication No. 12-E014-EF.
20. Hussey PS, Mulcahy AW, Schnyer C, Schneider EC. Closing the quality gap: revisiting the state of the science (vol 1: bundled payment: effects on health care spending and quality). Evidence

REVIEW ARTICLE

Evaluating the State of Quality-Improvement Science through Evidence Synthesis: Insights from the Closing the Quality Gap Series

reports/technology assessments, no. 208.1. Rockville, MD: Agency for Healthcare Research and Quality; 2012 Aug. AHRQ Publication No. 12-E007-EF.

21. Mauger Rothenberg B, Marbella A, Pines E, Chopra R, Black ER, Aronson N. Closing the quality gap: revisiting the state of the science (vol 6: prevention of healthcare-associated infections). Evidence reports/technology assessments, no. 208.6. Rockville, MD: Agency for Healthcare Research and Quality; 2012 Nov. AHRQ Publication No. 12-E012-EF.

22. McPheeters ML, Kripalani S, Peterson NB, et al. Closing the quality gap: revisiting the state of the science (vol 3: quality improvement interventions to address health disparities). Evidence reports/technology assessments, no. 208.3. Rockville, MD: Agency for Healthcare Research and Quality; August 2012. AHRQ Publication No. 12-E009-EF.

23. Totten AM, Wagner J, Tiwari A, O'Haire C, Griffin J, Walker M. Closing the quality gap: revisiting the state of the science (vol 5: public reporting as a quality improvement strategy). Evidence reports/technology assessments, no. 208.5. Rockville, MD: Agency for Healthcare Research and Quality; 2012 Jul. AHRQ Publication No. 12-E011-EF.

24. Viswanathan M, Golin CE, Jones CD, et al. Closing the quality gap: revisiting the state of the science (vol 4: medication adherence interventions: comparative effectiveness). Evidence reports/technology assessments, no. 208.4. Rockville, MD: Agency for Healthcare Research and Quality; 2012 Sep. AHRQ Publication No. 12-E010-EF.

25. Williams JW, Jackson GL, Powers BJ, et al. Closing the quality gap: revisiting the state of the science (vol 2: the patient-centered medical home). Evidence reports/technology assessments, no. 208.2. Rockville, MD: Agency for Healthcare Research and Quality; 2012 Jul. AHRQ Publication No. 12-E008-EF.

26. Shekelle PG, Pronovost PJ, Wachter RM, et al; PSP Technical Expert Panel. Assessing the evidence for context-sensitive effectiveness and safety of patient safety practices: developing criteria. Rockville, MD: Agency for Healthcare Research and Quality; 2010 Dec. AHRQ Publication No. 11-0006-EF.

27. Shekelle PG, Pronovost PJ, Wachter RM, et al. Advancing the science of patient safety. Ann Intern Med 2011 May 17;154(10):693-6. DOI: http://dx.doi.org/10.1059/0003-4819-154-10-201105170-00011

28. Foy R, Ovretveit J, Shekelle PG, et al. The role of theory in research to develop and evaluate the implementation of patient safety practices. BMJ Qual Saf 2011 May;20(5):453-9. DOI: http://dx.doi.org/10.1136/bmjqs.2010.047993

29. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 7;4:50. DOI: http://dx.doi.org/10.1186/1748-5908-4-50

30. McDonald K. Considering context in quality improvement interventions and implementation: concepts, frameworks, and application. Academic Pediatrics. (Forthcoming).

31. CHIPRA pediatric quality measures program [Web page on the Internet]. Rockville, MD: Agency for Healthcare Research and Quality; updated 2013 Mar [cited 2013 July 6. Available from: www.ahrq.gov/policymakers/chipra/pqmpback.html.

32. National advisory council subcommittee: identyfing health care quality measures for medicaid-eligible adults: background report [Web page on the Internet]. Rockville, MD: Agency for Healthcare Research and Quality; updated 2011 Dec [cited 2012 Sep 12]. Available from: www.ahrq.gov/legacy/about/nacqm11/.

33. News release: HHS awards ARRA funds to establish a center of excellence in research on disability services, care coordination and integration [press release on the Internet]. Washington, DC: US Department of Health and Human Services; 2011 May 7 [cited 2012 Sep 24]. Available from: www.hhs.gov/news/press/2010pres/05/20100506a.html.

34. 2012 annual progress report to congress: national strategy for quality improvement in health care [monograph on the Internet]. Washington, DC: US Department of Health and Human Services; 2012 Apr [corrected 2012 Aug; cited 2013 May 23]. Available from: www.ahrq.gov/workingforquality/nqs/nqs2012annlrpt.pdf.

35. Measure applications partnership [Web page on the Internet]. Washington, DC: The National Quality Forum; c2013 [cited 2013 Jan 17]. Available from: http://qualityforum.org/map/.

Toward Improvement

'What is everyone learning?' Asking the question that way will help clinicians and researchers see further in navigating toward improvement.

— Donald Berwick, MD, b 1946, former Administrator of the Centers for Medicare and Medicaid Services and former President and Chief Executive Officer of the Institute for Healthcare Improvement