Medical Care
Research and Review
Volume XX Number X
Month XXXX xx-xx

10.1177/1077558708330424
http://mcr.sagepub.com
hosted at
http://online.sagepub.com

What Can We Learn From Quality Improvement Research?

A Critical Review of Research Methods

Jeffrey A. Alexander
Larry R. Hearld
University of Michigan

This article presents a systematic review of the research methods used to study quality improvement (QI) effectiveness in health care organizations. The review relied on existing literature as well as emergent themes to identify types of QI programs (e.g., data/feedback, information technology, staff education) and quality outcomes (e.g., mortality, morbidity, unnecessary variation). Studies were separated into four categories according to the type of organization in which the QI program was introduced: (a) hospital, (b) nursing home, (c) physician group, and (d) other health care organization. Results of the review indicate that most QI effectiveness research is conducted in hospital settings, is focused on multiple QI interventions, and utilizes process measures as outcomes. The review also yielded substantial variation with respect to the study designs used to examine QI effectiveness. The article concludes with a critique of these designs and suggestions for ways future research could address these shortcomings.

Keywords: *quality improvement effectiveness; literature review; health care organizations*

Significant opportunities exist for improvement in the clinical quality of care provided by health care organizations. Existing evidence, for example, suggests that one fourth of hospital deaths may be preventable; 180,000 people may die each year, partly as a result of iatrogenic treatments. Moreover, one third of some clinical procedures may expose patients to risks without improving their health; one third of drugs prescribed may not be indicated; and one third of laboratory tests showing abnormal results may not be followed up by clinicians (Brook at al., 1990; Dubois & Brook, 1988; Institute of Medicine, 2000; Leape, 1994).

In response to these issues, the Institute of Medicine's seminal report on the state of health care quality in the U.S. health care system (Institute of Medicine, 2000) and

Authors' Note: This article, submitted to *Medical Care Research and Review* on April 10, 2008 was revised and accepted for publication on October 31, 2008.

This research was funded by the Forum on Quality Improvement Research and Implementation, Institute of Medicine, National Academies of Science.

its subsequent calls for system redesign (Institute of Medicine, 2001) have precipitated an increase in quality improvement (QI) efforts in hospitals and other health care delivery organizations. QI posits that the quality of goods and services depends foremost on the processes by which they are designed and delivered. Hence, QI focuses on understanding, controlling, and improving work processes rather than on correcting individuals' mistakes after the fact (Deming, 1986; Ishikawa, 1985; James, 1989; Walton, 1990). QI also emphasizes analyzing the root causes of problems, taking appropriate steps to make work processes predictable, and then continuously improving process performance (Juran, 1989). In health care settings, QI varies considerably in its operational focus, targeted outcomes, and level of provider involvement. In the most general sense, health care QI involves any type of planned organizational change or intervention designed to improve some aspect of quality of care.

Early on, health care organizations focused their QI efforts on administrative challenges, such as billing, scheduling, recordkeeping, and related management functions (Health Care Advisory Board, 1992). By the mid-1990s, hospitals and other delivery organizations began applying QI to clinical processes and reported success in lowering costs of care, lengths of stay, and patient charges without adversely affecting patient outcomes for specific procedures and conditions (Brothers, Robison, & Elliott, 1997; Ellrodt, Conner, Rieidinger, & Weingarten, 1995; Gregor, Pope, Werry, & Dodek, 1996). Some also reported success in reducing adverse drug events, postoperative complications, hospital-acquired infections, and other quality problems now recognized as indicators of patient safety (Barnes, Lawton, & Briggs, 1994; Carey & Teeters, 1995; Civetta, Keller, & Wennberg, 1996). Although most of these reports are anecdotal, in the past decade a significant body of empirical research has been published on the effects of QI interventions in health delivery organizations.

Given the development of QI effectiveness research, it is reasonable to ask whether such research can provide managers and policy makers with evidence about how well such efforts are working, especially with respect to their impact on quality-related outcomes. However, research on the effectiveness of QI efforts has not been systematically examined in such a way as to draw reliable conclusions about what works, when it works, or how it works (Shortell, Bennett, & Byck, 1998). One challenge in this regard is the difficulty in comparing studies because of significant differences in interventions, study settings, research designs, measurement, and data sources. A second challenge stems from methodological and conceptual problems in the studies themselves that affect both the internal validity of the findings and their generalizability.

The purpose of this article is to critically review the methods employed in empirical studies of QI change in health care organizations. Specifically we will (a) classify and describe the extant research on QI effectiveness, (b) critically assess the methods employed in QI effectiveness research, and (c) offer recommendations about how QI studies might be improved to enhance the validity, comparability, and generalizability of the findings of QI studies. The article concludes by providing a

set of caveats for the application of extant research findings to improving organizational practices in real-world settings.

New Contribution

This review extends previous QI research by examining methodological and conceptual problems in a broad and growing body of QI literature. Our review indicates significant variation in the quality outcomes, interventions, and methods used to study QI effectiveness, as well as a number of potential methodological problems with many of the studies. Both these types of issues raise important questions about research quality and the conclusions that can be drawn from this body of literature. The study will use these critiques as the basis for a series of conceptual and methodological recommendations to improve QI research. Improved methods in QI research are likely to result in enhanced internal validity and generalizability of research findings and, perhaps as important, identification of QI interventions that hold the most promise for improving quality of care.

Analytic Framework

For purposes of this review, health care QI is defined as actions for improving the processes and outcomes of health care, including increasing value; improving responsiveness to customers and consumers; improving outcomes in the areas of safety, effectiveness, timeliness, patient centeredness, equity, and efficiency; reducing variation in outcomes; and increasing organizational adoption and implementation of continuous improvement methods in ongoing operations. The review uses the term *intervention*(*s*) to refer to the broad set of changes that are included under this definition of QI. Although QI practitioners have not always incorporated consistent terminology when engaging in these activities, we believe that the adoption of a precise definition and a common term to represent the set of QI activities is important to facilitate a clear and cohesive discussion of a broad and diverse literature.

The review relied on existing literature as well as emergent themes to identify types of QI programs (e.g., data/feedback, information technology, staff education). A second dimension of our review was an assessment of QI effectiveness studies by organization type. Studies were separated into four categories according to the type of organization in which the QI program was introduced: (a) hospital, (b) nursing home, (c) physician group, and (d) other health care organization.

The analytic framework adopted for this review addresses several characteristics of QI programs and the research on such programs. First, classifying the QI interventions into categories narrows the focus to more parsimonious, common groupings to allow comparisons across types of QI changes. Given both the proliferation of research in this area and the idiosyncratic quality of QI efforts, the use of common

frameworks is important for comparing study results while still allowing organizations and researchers to test the effects of QI programs that accommodate local conditions. Second, the framework allows comparisons across different types of organizations and allows us to evaluate whether different types of QI studies are more or less common to different types of organizations. Differences may emerge for several reasons, such as resource availability or even applicability of certain types of QI changes in different settings. For example, hospitals may be able to provide greater financial support than physician practices to implement and sustain advanced information technology and sophisticated data feedback systems, thus giving rise to more research in this area.

Method

Inclusion–Exclusion Criteria

Searches were limited to peer-reviewed articles in English-language journals and were confined to original, empirical research on health care organizations, including hospitals, physician groups, nursing homes, and other general health care providers. Non–health care organizations were excluded from analysis, as well as behavioral health care organizations (mental health, substance abuse treatment). Because our review was limited to original, empirical research, we excluded theoretical or conceptual articles on QI models, meta-analyses, and opinion or editorial articles. However, qualitative studies of QI effects were included for review. Given our objective of assessing purposive efforts by organizations to *improve* quality, we excluded purely descriptive studies of QI programs and studies that examined associations between general organizational characteristics and quality. For example, studies examining differences in quality across different nursing home ownership types did not qualify for review.

Although the searches included studies across all years, the review concentrated on research published since 2000. Therefore, although we include studies prior to 2000, we cannot make claims to exhaustiveness.

Search strategy. MEDLINE, CINAHL, and Evidence-Based Medicine Reviews were the primary databases used to identify eligible studies. Together these databases capture the majority of the extant health services research on QI. Primary keyword searches of *effect of quality improvement* were done in these databases. These initial searches returned 7,842 articles. Given the extensive number of articles returned by this initial search, we combined predictor and outcome variables as search terms to delimit the body of literature for review. A secondary search that used the outcome variables as keywords in combination with the primary search

term was used to narrow our search. For example, for outcomes we queried *effect of quality improvement* together with *mortality*. In total, nine outcome variables were paired with each predictor, including *mortality*, *morbidity*, *unnecessary variation*, *overuse*, *underuse*, *misuse*, *errors*, *patient safety*, and *patient satisfaction*. Keywords were mapped to their respective Medical Subject Headings to broaden our search and ensure we captured the corpus of literature relevant to our review. A final search included the type of organization as a third keyword. Using the same example, the final search would include *effect of quality improvement*, *mortality*, and *hospitals* all in the same search. Three specific search terms related to organizational type were included as keywords: *hospital*, *nursing home*, and *physician practice*. These keywords were also mapped to their respective Medical Subject Headings.

Because our focus also included more general organizational changes to improve quality, we repeated our searches using *organization change* as a search term in place of *quality improvement*. For example, the search was modified to search for articles with *effect of organizational change*, *mortality*, and *hospitals* all in the text. A supplementary search strategy included bibliography reviews of relevant articles to identify additional articles for analysis.

Full articles were reviewed for variables when abstracts were not available or when the abstract did not contain enough information to accurately determine whether an article was within scope. Subsequent reviews of article bibliographies contained within the secondary database brought the number of fully reviewed articles to 185. They are listed in the appendix.

Coding and synthesis of literature. Prior to initiating full review, the authors reviewed the same 10 articles to assure reliability and accuracy of article coding. All articles were then reviewed to identify study design characteristics, organizational characteristics, predictor and outcome variables, and the direction and significance of the relationships examined. A standardized form was used to ensure consistent coding across articles and reviewers. Study design characteristics included sample size, whether the study was observational or a randomized controlled trial, whether the study was cross-sectional or longitudinal, and whether the study included a comparison or control group. Organizational characteristics included the type of organization (e.g., hospital, nursing home), ownership (e.g., for profit), and the organizational unit where the study was conducted (e.g., intensive care unit), when identified by the study. Predictor variables focused on the types of QI interventions studied, such as benchmarking or the use of clinical reminders. Outcome variables included quality outcomes such as mortality, morbidity, and patient satisfaction.

Because we were interested in QI and process change as general phenomena, the review does not focus on specific QI approaches (e.g., Continuous QI; Plan, Do, Check, Act; Six Sigma). Instead, we chose to focus on the granular interventions that make up these general approaches to process change. The focus on discrete

interventions and the use of consistent definitions is important for creating consistency and enabling synthesis across a diverse and largely noncumulative literature. Therefore, a second analytic step focused on grouping the QI interventions into one of nine categories: (a) data/feedback/reminders, (b) case/disease management, (c) evidence-based medicine, (d) information technology, (e) task integration/decomposition, (f) multidisciplinary teams, (g) practitioner education, (h) patient education, and (i) multiple interventions (Grimshaw et al., 2004; McLaughlin & Kaluzny, 2006). The *data/feedback/reminders* category was defined as the distribution of summary performance information or information to guide behavior of practitioners. *Case/disease management* was defined as a formal management program focused on managing the complex health care needs of a patient or a population of patients with a specific clinical condition. *Evidence-based medicine* was defined as the development or implementation of evidence-based practices in the delivery of care and included practices such as the use of critical pathways or care guidelines. *Information technology* was defined as the use of information technology in the delivery of care to improve quality. Examples of *information technology* used as QI interventions include the implementation of computerized physician order entry (CPOE) and the adoption of electronic medical records. *Task integration/decomposition* was defined as the development of new work processes or the redesign of existing work processes through the consolidation or disaggregation of work activities and behaviors. Examples of activities that fell within this category include introduction of restrictive policies relating to drugs commonly associated with adverse events, addition of steps to the discharge planning process to improve follow-up, and changes in the physical design of the work space to facilitate hand washing. *Multidisciplinary teams* involved the creation of a team of health professionals across different disciplines or additions of new members to an existing team working to care for patients. *Practitioner education* focused on the dissemination of educational materials and clinical recommendations to practitioners, either personally or via mass communication, while *patient education* included the dissemination of educational materials and recommendations to patients. Finally, the *multiple intervention* category was created to capture those studies in which more than one intervention was examined but the effects of the individual interventions could not be discerned from the article. In virtually all cases, these studies were designed to capture the combined effects of multiple QI interventions. For example, a study examining the combined effects of the implementation of clinical practice guidelines and the education of practitioners on the use of these guidelines would be classified as multiple interventions. If a study examined more than one type of QI intervention and clearly distinguished the effects of each intervention, the QI interventions were assigned separately to their respective categories. For instance, studies examining the independent effects of clinical practice guidelines and the physician educational seminars on the use of these guidelines were assigned to the evidence-based medicine and practitioner education categories.

Similar coding was done to create consistency across categories of outcome variables. In total, eight categories were used to classify outcome variables. Outcome categories generally corresponded with our search term strategy (e.g., mortality, patient satisfaction). Errors and patient safety were collapsed into one category for analytical purposes. An additional category was created to reflect those studies that included process measures as outcomes. *Process measures* were defined as services or activities done to or on behalf of a patient that are assumed to contribute to a positive health outcome. Examples of process measures include the number or rate of recommended procedures done for patients (e.g., lab tests for diabetic patients, adherence to asthma guidelines) and increased adherence to quality protocols (e.g., improved patient assessment and documentation).

A final analytic step involved assigning a 0 or 1 value to each variable (e.g., organizational type, outcome) to indicate whether the study included that variable. For example, a study examining the association between the introduction of an information technology system and medical errors in a hospital would have information technology and errors and hospital coded as 1. These values were subsequently used to calculate variable frequencies by summing across reviewed studies. For articles including more than one predictor or outcome of interest, each variable was assigned a 0 or 1 value. For example, a study examining the effect of information technology and process redesign on medical errors would assign a value of 1 to both information technology and process redesign.

An important consideration in interpreting the results reported in this review is the tendency of the health care literature not to publish findings that reflect null results. Journals are disposed toward those studies that demonstrate positive changes or dramatic findings, or authors are reluctant to submit their work for publication of "no difference" as the main finding. Although most health services researchers would probably take the position that null findings are as informative as those that are statistically significant, this is not the reality of publishing. Although it is difficult to estimate the degree of bias associated with unpublished QI studies that find null results, a generally conservative position would suggest that many studies do not appear in published form because they are not able to demonstrate results that differ significantly from the null. From a different perspective, there is likely to be considerable QI activity (successful and unsuccessful) in health delivery organizations that never reaches the peer-reviewed literature. These activities likely originate from attempts to solve quality or operational problems in delivery organizations without regard to meeting the standards of internal or external validity required by research studies. Dissemination of these practices occurs, if at all, in the practitioner literature or by word of mouth or other informal means. If these assumptions are accepted, results reported in the literature are most likely to both overstate and understate the degree to which QI efforts result in improvements in quality outcomes.

Results

QI Study Settings, Interventions, and Outcomes

For this study 185 articles relating to the effects of QI were fully reviewed. The majority of articles reviewed (70%) were published since 2000, and nearly 90% were published in the past decade. The range of journals used to publish these studies reflects the growth and diversity in this literature. Our review found studies in general clinical and health services journals as well as domain-specific clinical journals such as those in nursing and pharmacy.

QI research is not evenly distributed across all organizational settings. Nearly 62% (114 articles) focused on QI interventions in hospitals (Table 1). It is noteworthy that among the hospital-based QI investigations, the majority have been conducted in university teaching hospitals, and a relatively small proportion in community hospitals. Although it is understandable that the university hospitals are more conducive settings to research on the effects of QI interventions, the predominance of these hospitals as the context for QI raises issues about the applicability of study findings to other settings, notably nonuniversity community hospitals. This issue is particularly important given that most hospital admissions occur among community hospitals rather than in teaching hospitals.

In comparison with hospital-based studies, 24% and 11% of reviewed studies were conducted in physician practices and other health care organizations, respectively. A surprisingly low number of nursing home studies (3.8%) explicitly examined the relationship between formal QI interventions or changes and quality-related outcomes. A close review of the studies identified in our searches suggests that much of the quality-related research in nursing homes has focused on correlational relationships between general organizational characteristics and quality, with relatively little attention paid to interventions or organizational changes that may result in improved quality. An example of such out-of-scope research is a study by Dellefield (2006) that examined whether the prevalence of pressure ulcers was associated with nursing staff hours, nurse specialization, ratio of nurse supervisors to frontline staff, or for-profit ownership. Studies such as this one are helpful in identifying quality deficiencies but stop short of testing targeted changes to improve these deficiencies.

Similar variation was observed with respect to the types of QI interventions examined in the studies (Table 1). More than 30% of all reviewed studies focused on multiple interventions (e.g., physician education and evidenced-based medicine, together with case disease management), and the effects of the interventions or intervention components could not be independently identified. This was particularly true in QI studies set in hospitals and physician practice settings.

These multi-intervention studies were followed in frequency by practitioner education (21.1%) and the use of data and feedback interventions (16.2%). At the lower degree of frequency are those QI studies that focused on the effects of information

Table 1
Number of Reviewed Articles by Category of Quality Improvement and Organization Type

	Organization Type									
Type of Quality Improvement Intervention	Total	% of Total Articles	Hospital	% of Hospital Articles	Nursing Home	% of NH Articles	Physician Practice	% of Physician Practice Articles	Other	% of other Hospital Articles
Data/feedback	30	16.2	18	15.8	0	0.0	10	22.7	2	10.0
Case/disease management	8	4.3	5	4.4	1	14.3	1	2.3	1	5.0
Evidence based medicine	16	8.6	13	11.4	1	14.3	2	4.5	0	0.0
Information technology	7	3.8	5	4.4	1	14.3	1	2.3	0	0.0
Task integration/ decomposition	21	11.4	14	12.3	2	28.6	3	6.8	2	10.0
Multidisciplinary teams	14	7.6	11	9.6	1	14.3	2	4.5	0	0.0
Practitioner education	39	21.1	19	16.7	1	14.3	10	22.7	9	45.0
Patient education	7	3.8	2	1.8	0	0.0	3	6.8	3	15.0
Multiple interventions	56	30.3	35	30.7	1	14.3	14	31.8	6	30.0
Number of articles	185		114		7		44		20	
% of total articles			61.6		3.8		23.8		10.8	

technology (3.8%) and patient education (also 3.8%). While these numbers are perhaps surprising in light of recent emphasis on information technology and consumer engagement in health care, a review of the studies that did not meet our criteria provide some explanations for these low numbers. Similar to nursing home studies, much of the literature in these areas has thus far focused on describing the development or implementation of information technology or consumer engagement, with considerably less attention paid to evaluations of how these changes relate to QI. Some of this tendency may be explained by the nature of the interventions themselves. information technology interventions (such as electronic medical records), for example, are meta-interventions, with hundreds or even thousands of specific features potentially affecting quality. These characteristics produce a virtually infinite number of combinations of features and local environmental circumstances, making study designs particularly challenging.

Process measures were the most common outcome measures used in the reviewed studies, with nearly 45% of all reviewed studies examining the effects of QI interventions on processes of care (Table 2). While process measures were employed in QI studies in all organizational settings, studies in physician practices were particularly oriented toward process measures, with more than 65% of reviewed articles in physician practice settings using such measures as dependent variables. In contrast, only slightly more than one third of reviewed hospital studies used process measures as outcomes. Employing process measures that are purportedly related to clinical QI, versus direct measures of clinical QI, may be a function of the relative accessibility of such measures or their relative importance in the context of specific QI interventions (e.g., beta blockers, particular forms of evidenced-based medicine).

Morbidity (30.8%) was the second-most-common outcome used across all QI studies and was the most frequently used outcome in nursing home studies (57.1%). Underuse (18.4%), overuse (17.3%), and unnecessary variation (17.3%) ranked third, fourth, and fifth as the most frequently used outcomes across all settings. Errors/patient safety (6.5%) and patient satisfaction (6.5%) were tied for the least frequently studied outcomes in studies that evaluated the effectiveness of formal QI interventions or changes.

Data Sources and Collection Methods

A wide variety of data collection methods were observed in the QI studies reviewed (Table 2). These ranged from the use of administrative databases to chart abstraction, survey methods, or even direct observation or surveillance of providers. Of the methods employed, chart abstraction, use of administrative databases, survey methods, and multiple data collection approaches were the most commonly used, representing 38%, 37%, 29%, and 23% of all studies reviewed, respectively. The relative use of these methods appeared to vary somewhat by the organizational context in which the study was conducted. For example, whereas chart abstraction

Table 2
Number of Reviewed Articles by Outcome and Organization Type

	Organization Type									
Outcome[a]	Hospital	% of Hospital Articles	Nursing Home	% of Nursing Home Articles	Physician Practice	% of Physician Practice Articles	Other	% of Other Articles	Total	% of Total Articles
Mortality	15	13.2	2	28.6	0	0.0	0	0.0	17	9.2
Morbidity	40	35.1	4	57.1	10	22.7	3	15.0	57	30.8
Unnecessary variation	23	20.2	1	14.3	6	13.6	2	10.0	32	17.3
Overuse	14	12.3	0	0.0	12	27.3	6	30.0	32	17.3
Underuse	23	20.2	1	14.3	5	11.4	5	25.0	34	18.4
Errors/patient safety	9	7.9	0	0.0	2	4.5	1	5.0	12	6.5
Patient satisfaction	8	7.0	1	14.3	3	6.8	0	0.0	12	6.5
Process	40	35.1	3	42.9	29	65.9	11	55.0	83	44.9
Data collection method[a]										
Administrative/ clinical database	38	33.3	5	71.4	17	38.6	9	45.0	69	37.3
Chart abstraction	50	43.9	4	57.1	13	29.5	4	20.0	71	38.4
Focus group	1	0.9	0	0.0	0	0.0	1	5.0	2	1.1
Diary/log books	0	0.0	0	0.0	2	4.5	0	0.0	2	1.1
Interviews	1	0.9	0	0.0	0	0.0	1	5.0	2	1.1
Process mapping	1	0.9	0	0.0	0	0.0	0	0.0	1	0.5
Simulated patient encounters	0	0.0	0	0.0	0	0.0	1	5.0	1	0.5
Survey	26	22.8	2	28.6	19	43.2	8	40.0	55	29.7
Surveillance/observation	6	5.3	1	14.3	0	0.0	0	0.0	7	3.8
Survey/patient assessment	3	2.6	0	0.0	2	4.5	1	5.0	6	3.2
Not stated	1	0.9	0	0.0	0	0.0	0	0.0	1	0.5
Multiple data collection methods	21	18.4	5	71.4	10	22.7	6	30.0	42	22.7
Number of articles reviewed	114		7		44		20		185	
% of total articles	61.6		3.8		23.8		10.8			

a. Summing across outcomes and data collection methods within an organization type category does not result in the total number of articles because of articles including multiple outcomes and data collection methods.

appeared to be predominant in hospital-based QI studies, surveys and administrative databases tended to be more common in QI studies focusing on physician groups or clinical practices.

The wide variety of data collection methods used to assess outcomes in QI studies may be viewed as a contributing factor to the lack of comparability of study findings. However, another perspective suggests that such variety is necessary, given the types of interventions being studied or the outcomes of relevance in particular organizational contexts. For example, surveillance and observation of clinical behavior may be more relevant than a review of patient charts if the intervention is focused on changing practice behavior or patterns of interactions between patients and physicians. Alternatively, multiple data collection methods might be best suited for studies that examine not only outcomes but the process of implementing a QI intervention.

It is also worth noting that the source of data for outcomes in QI studies is subject to considerable controversy and may affect comparability of the study results. For example, physicians often object to constructing quality-of-care outcome measures from discharge abstracts or claims data, asserting that such measures are too gross to truly capture the quality of care provided by practitioners. These groups tend to reject almost categorically studies in which outcomes are based on administrative databases. On the other hand, researchers often criticize the expense associated with pulling patient charts and the restrictions such techniques place on obtaining adequate sample sizes to generalize to organizational or patient populations. Although such differences are not likely to be resolved anytime soon, those charged with evaluating QI effectiveness studies should carefully weigh the trade-offs between obtaining large samples of organizations and patients, on one hand, and the more nuanced clinical outcome data obtainable from patient records.

Our review also examined sample sizes of both organizations and participants included in QI studies. Organization sample size indicates the number of organizational settings where the QI intervention was studied. For example, a study by Chu et al. (2003) examined the effect of external feedback on care improvement processes related to pneumonia across 36 community hospitals. The organization sample size for this study was 36 hospitals. In contrast, the participant sample focuses on the number of patients or practitioners included in a study. In the Chu et al. study, the participant sample size was 2,057 patients. In this case, the researchers examined the medical charts of 2,057 patients treated within the 36 study hospitals. A distinction between organization and participant samples is important for highlighting the nested and contextual nature of QI interventions. Ignoring such distinctions could lead to inappropriate comparisons across studies, which in turn could lead to incorrect conclusions regarding the effects of a particular QI intervention.

Table 3 indicates that the median organizational sample size for all QI studies reviewed was 2, indicating that most of these studies would not be large enough to test statistically for differences in either QI effects or implementation effectiveness across organizational conditions. Both the median and modal sample size for hospital QI

Table 3
Sample Size Descriptives by Organization Type and Sample Type

	Hospital		Nursing Home		Physician Practice		Other		Total	
	Organizations	Participants	Organizations	Participants	Organizations	Participants	Organizations	Participants	Organizations	Participants
Mean	82	22,054	37	1,464	66	23,961	58	30,174	75	22,802
Median	1	526	19	251	42	609	15	947	2	565
Standard deviation	340	107,728	43	3,018	80	86,410	82	115,748	275	100,326
Maximum	1,784	911,407	133	7,623	374	455,843	270	464,200	1,784	911,407
Minimum	1	4	10	137	1	15	1	50	1	4
Number of articles	114	114	7	7	44	44	20	20	185	185

studies is one hospital, with a mean sample size of 82 hospitals. Mean hospital sample size is inflated by a small number of studies with very large samples. By contrast, the median sample sizes for QI studies in nursing homes, physician practices, and other health delivery organizations are larger, with 19, 42, and 15 organizations, respectively. Although these samples are all larger than those of most hospital QI studies, most are convenience samples and therefore may not represent a defined population of organizations. On the positive side, they are frequently large enough to test statistically for differential effects of QI changes across a limited number of organizational conditions.

Similar to organizational samples, descriptive data for the participant sample sizes reveal that very large mean values appear to be driven by a small number of studies with participant samples in the multiple thousands. These studies often rely on discharge abstracts or Medicare data to construct quality measures. For example, a study by Ferguson et al. (2003) included 267,917 patients undergoing coronary artery bypass graft surgery; it abstracted process data from hospital discharge data sets. Median values of participant sample size suggest that the more typical participant sample size averages in the several hundreds.

Study Designs

Table 4 presents design characteristics of the studies reviewed. The majority (62%) of QI effectiveness studies employed observational techniques, and 38% utilized experimental/randomized control trial (RCT) designs. Whereas RCT studies have long been considered the gold standard for purposes of establishing internal validity and causality, these designs may not be practical in an organizational context. For example, it may be difficult to randomize patients or providers to experimental and control groups if there is the potential for withholding needed treatment to these groups. The use of RCTs may also be impeded by "ownership" of the QI intervention by key stakeholders, commitment by leadership to the change, and other local conditions that are difficult to reconcile with the randomization and control features associated with such designs. Further, these designs may be subject to the effects of contamination if pilot sites used for testing the intervention are not sufficiently insulated from other units within an organization or practice. These designs also rely heavily on self-selected participants, who may not be representative of the targeted population for the intervention or may be predisposed to be enthusiastic proponents of a particular QI technique. The final disadvantage of RCTs is that they are conducted under, as the name implies, highly controlled conditions. These conditions may not reflect the reality of the context in which the QI practice might normally be introduced and used. Hence, this context-free form of QI effectiveness research may place severe limitations on judging whether the intervention would likely be as effective in other contexts or situations. Given the well-documented difficulties in implementing and diffusing best practices in health care delivery, lack of attention to context in RCTs would seem to be antithetical to the goal of greater uptake of best practices in QI.

Table 4
Number of Reviewed Articles by Design Characteristics

Characteristics	Count	% of Total[a]
Randomized control trial (RCT) vs. observational		
RCT	71	38.4
Observational	114	61.6
Cross-sectional vs. longitudinal		
Cross-sectional	91	49.2
Longitudinal	94	50.8
Study use comparison group?		
Yes	129	69.7
No	56	30.3
Study measure implementation?		
Yes	14	7.6
No	171	92.4
Results discuss effect size?		
Yes	137	74.1
No	48	25.9
Study includes cost–benefit analysis?		
Yes	23	12.4
No	162	87.6
Study use qualitative or mixed methods?		
Yes	7	3.8
No	178	96.2
Unit of analysis		
Patient	84	45.4
Physician	14	7.6
Staff	10	5.4
Team	1	0.5
Unit/department	18	9.7
Organization	46	24.9
Region/network	3	1.6
Multiple units	9	4.9

a. % of total represents total sample of articles reviewed ($N = 185$).

Although observational studies tend to dominate in QI effectiveness research, this may be more a function of the constraints that RCTs carry rather than the preferred approach to the conduct of such studies. We observed, for example, that many QI studies, whether conducted in hospitals, physician practices or other health care settings, were studies of opportunity, that is, unfunded studies of an administratively or clinically mandated change in an organizational practice. Often these opportunistic studies do not allow sufficient time for careful design or even selection of adequate control groups. For example, fully 30% of the QI studies reviewed contained no comparison group, thus making it difficult to determine

whether the observed outcomes associated with the QI change were equally likely to occur in populations or organizations that did not experience this change.

In a similar vein, only about half of all QI studies reviewed assessed quality-related outcomes with longitudinal data. This suggests that many QI studies are correlational and do not satisfy the temporal sequencing required to infer causation between the QI intervention and the outcome of interest. Further, the 94 studies that did employ longitudinal methods revealed considerable variability in the way temporal data were incorporated in the study design. For example, relatively few of these studies employed multiple time series observations of the dependent variable. This suggests that these investigations do not assess temporal patterns of change in the outcome of interest. Employing multiple time series observations of dependent variables would permit investigators to assess whether changes in outcomes were instantaneous or alternatively required a substantial time period to realize. Similarly, multiple time series observations would enable detection of the presence of change trends prior to the introduction of the QI intervention.

Longitudinal QI studies also differ significantly in their duration. The median length of time during which the dependent variable was assessed was less than 1 year, with many significantly shorter than this. The average study length would be reduced even further if cross-sectional studies were incorporated into this calculation. This may suggest that many studies are of insufficient length to determine the long-term sustainability of quality changes that might result from the QI intervention.

Relatively few of the studies reviewed formally assessed and measured the implementation of the intervention being examined. Only about 7.5%, or 14 of the 185 studies, assessed the implementation of the intervention under investigation. This raises a number of issues pertaining to QI effectiveness. For example, without a formal assessment of intervention implementation, it is unclear whether the findings of no effect are due to the lack of efficacy of the intervention, on one hand, or inappropriate or incomplete implementation of the intervention, on the other.

Similarly, relatively few of the reviewed QI effectiveness studies contained any type of cost–benefit or cost-effectiveness analysis. Only 12% of all studies contained such analysis. This suggests that most QI studies considered only the relationship between the QI change or intervention on one hand, and the outcome of interest on the other. This may not be enough information to inform managers and policy makers of the significance of the intervention or its appropriateness for broader application or dissemination. In other words, it is not sufficient to know whether there is a significant relationship between the introduction of a QI change and an outcome of interest, say, mortality or decrease in the use of undesirable services. Even if the effect size is relatively large in such cases, without knowing the cost of introducing the intervention relative to the methods employed, managers are in a weak position to determine the value of a particular QI approach. Indeed, even studies that purport to include a cost–benefit analysis often do not actually measure the cost of adopting or implementing an intervention but simply focus on cost savings that result from

these changes. Conversely, studies sometimes focus on demonstrating the cost of the intervention but do not attempt to quantify the benefits or savings that result from that intervention.

Although not the primary focus of this article, our review indicates that correlations obtain between study design characteristics used in QI research, on one hand, and reported outcomes, on the other. For example, 38% of the reviewed articles employed an RCT design (vs. an observational design). These studies, however, were substantially less likely to report significant improvements in quality than were the observational studies (51% vs. 68%, respectively). This suggests that more rigorous designs may be better able to control for unobserved factors other than the QI intervention that can influence change in quality outcomes. Similar, although smaller, differences were observed for studies conducted in single versus multiple organizations. In all, 45% of reviewed articles were conducted in single organizations. These studies reported significant changes in 67% of relationships examined versus only 56% of relationships examined in multiple-organization studies. Finally, our review suggests that studies that employed a longitudinal design (vs. cross-sectional) reported substantially lower rates of positive quality outcomes. Only 56% of the QI relationships that used longitudinal analysis showed a positive association with quality outcomes, compared with almost 77% in cross-sectional studies. Again, these findings point to a conclusion that more rigorous designs are more likely to achieve findings of no difference when the relationship between QI change and quality outcomes is assessed. Put another way, to the extent that the literature is populated by weak study designs, the impact of QI changes may be overstated. Thus design factors may account for some of the inconsistency in findings and create additional difficulty in drawing clear conclusions about the effects of QI on quality-related outcomes.

QI Effectiveness Research: Critical Observations

A number of study design characteristics may limit the practical application of QI study findings beyond the organizations and participants involved in the studies. Among others, our review highlights issues of inadequate study duration, potential selection bias of study participants, and difficulty making generalizations to other organizations due to unique study contexts or variation in intervention characteristics. Another concern raised by this review is the limited attention given to the assessment of intervention implementation, which produces an incomplete picture of the complexity surrounding the contributions and sustained use of QI strategies.

One common methodological problem is short study duration. That is, the effects of most interventions reviewed were measured shortly after the introduction of the innovation or change (typically less than a year, sometimes less than a month), thus making it difficult to tell whether the results were sustainable over a protracted period. Organizational analysis would suggest that many changes have short half-lives and that organizations often revert to established routines or practices once the

stimulus of a new intervention or change has been removed. For example, a pre–post study by Szpunar, Williams, Dagroso, Enberg, and Chesney (2006) that examined the effects of automated guideline implementation on guideline adherence collected post data approximately 2 weeks after implementation. Given the tendency for organizational changes to experience both entropy and reversion to previous states, this is not a trivial matter. Short-term studies simply do not provide all the requisite information to ascertain whether the changes are going to have long-term effects on quality in these organizational settings.

From a different, but related, perspective, many complex QI interventions, such as those requiring simultaneous, multiple changes in organizational process and structure, may need extended periods to actually demonstrate intended results. A study by Schectman, Schroth, Verne, and Voss (2003) provides a good example of a complex, multifaceted intervention that required significant time to evaluate. This study examined the effects of guideline implementation, supported with practitioner education by opinion leaders and procedure audit and feedback, on variations in care for lower back pain. The intervention also included a patient education component. The evaluation period spanned more than 2 years. Indeed, the general organizational literature maintains that performance downturns often follow the introduction of major organizational change and positive results are observed only after such short-term disruptions are experienced and the new practices institutionalized in the organization. Both these issues point to the need for QI studies to take into consideration the long-term staying power of innovations or changes designed to improve quality.

Many QI studies incorporate either multiple interventions or interventions with multiple components (30% of reviewed studies). One interpretation is that multiple-intervention studies are a reflection of the systemic properties of QI and the emphasis on multifaceted strategies to improve quality. Such studies, however, present important challenges for researchers insofar as study designs must be able to assess not only how the combined effects of multiple interventions in combination affect outcomes but also, ideally, how, and the extent to which, individual components of the intervention contribute to these collective efforts. However, from a practical standpoint, parsing out the individual effects of multifaceted strategies is often not possible, and is perhaps even antithetical, to the notion of a systems-based approach to QI. That is, highly interdependent components of a complex system may prevent empirical separation because complex interactions between different components of such systems may produce synergistic results that are lost when the cumulative effect is disaggregated into its component parts. An example of tight interdependence between component parts that was commonly observed in the review related to studies of the effects of evidence-based medicine in which interventions typically required extensive educational interactions with practitioners to introduce the intervention and frequent reminders to improve adoption and implementation rates. For instance, a study by Harris et al. (2003) utilized an initial educational session to inform physicians of the appropriate protocols for prescribing antibiotics for acute respiratory infections, complemented by

exam room posters to encourage sustained use of the guidelines. Separating these components would likely undermine the use of the guideline intervention and thus preclude an evaluation of its effects. These difficulties notwithstanding, it is important, at a minimum, for researchers to assess the extent to which the individual components of a multifaceted QI intervention are implemented correctly and are operating according to the proposed theory of the intervention.

This is a common challenge faced in RCT studies, which often tend to treat the intervention as a single, undifferentiated process. For example, an RCT conducted by Horbar et al. (2004) examined the effect of a multifaceted intervention designed to promote evidence-based surfactant treatment for preterm infants. The intervention involved audit and feedback, evidence reviews, an interactive training workshop, and ongoing faculty support via conference calls and e-mail. Taken together, the components resulted in a significantly positive effect; however, the method of analysis did not allow a determination of how each component of the intervention operationally related to the others, the extent to which the program components were implemented successfully, or the contribution of the individual components of the program to the overall program effects.

In spite of (and because of) the practical limitations that often preclude a direct examination of the individual component effects for most studies, there is considerable value in research that can parse out the individual and combined effects of intervention components. Such studies, perhaps using factorial designs, could help empirically substantiate theoretical arguments about the systemic nature of QI interventions. A study by Eccles et al. (2001) provides a good example of how this may be done. To understand the effects of guideline use, audit and feedback, and reminder messages on radiology referrals, Eccles et al. used a 2×2 factorial design and randomized physicians to four groups: (a) received guideline only (control); (b) received audit and feedback; (c) received reminder messages; and (d) received audit, feedback, and reminder messages. The study found that the combined intervention of audit, feedback, and reminder messages, compared with the control group, significantly reduced the number of radiology referrals. It is interesting that when these component parts were separated, reminder messages significantly reduced the number of radiology referrals while audit and feedback did not have a significant effect on referral rates.

A related issue in QI effectiveness studies is that it is often difficult to determine whether the intervention was fully implemented. Indeed, the vast majority of studies assume that the change has been successfully implemented, and the research focuses on testing for an empirical association between the unmeasured innovation and the outcomes of interest. In multifaceted intervention designs, the complexity of such interventions makes the assessment of implementation particularly important because (a) the probability that all components of the intervention will not be implemented fully increases and (b) many multifaceted interventions assume that the components will function as a system to achieve the desired effects on quality. If one or more of

the elements of the system are not implemented, the system does not function as designed. The more complex the intervention, the more important it is to consider whether and how the individual components of the intervention combine to affect the systemic properties of the intervention. To these ends, there are increasing calls for alternative designs to explore questions related to QI effectiveness and implementation (Eccles, Grimshaw, Campbell, & Ramsay, 2003; Grol, Baker, & Moss, 2002). For example, process and implementation studies are needed to explore questions related to the adoption and sustainability of effective QI practices. The use of other methodological approaches, such as qualitative designs, holds the potential to improve our ability to address these issues. These approaches should be used in conjunction with more-traditional experimental and observational designs to address such key issues. Such multimethod approaches might also address questions related to why some RCTs did not achieve the anticipated effects on quality outcomes. Such explanations, for example, may point to whether the intervention itself was ineffective or the implementation of the intervention was incomplete.

Another issue with QI effectiveness studies is the influence of possible selection bias among those participating in these studies. Specifically, many QI effectiveness studies may suffer from problems of nonrepresentativeness of the physicians and patients who participated in such research, either because they were accustomed to doing these types of change or because they voluntarily participated in the study. In other words, it is possible that those involved in the change process, or being assessed as participants of the change, were not typical candidates for these changes. The most likely scenario is that only those motivated individuals or well-functioning groups might be likely to participate in QI studies, thus raising questions about whether a QI change strategy will work for all groups. For example, a study by Shojania et al. (1998) examined whether showing a computerized guideline in a CPOE system could reduce the use of vancomycin in a large, tertiary-care teaching hospital. While they observed a positive effect, the focus on users in a large teaching hospital, one with widespread use of CPOE, raises questions about generalization to hospitals without this capability or without previous acceptance of the value of CPOE. Many other QI effectiveness studies were also conducted in large university teaching hospitals, which have organizational cultures well-suited to innovation and change and where research is considered a primary activity. Such may not be the case for small community hospitals that have no teaching or research components to their mission.

Only a small proportion (12%) of the QI effectiveness studies reviewed examined the cost or potential costs of the interventions or conducted a cost-effectiveness assessment of the intervention. Typically, this was because of the lack of available, relevant financial information and also because the investigators apparently did not have the training or expertise to conduct cost-effectiveness assessments of the interventions studied. This is an important omission in the QI effectiveness literature because unless the value of the change to the organization is evaluated (not just its impact on quality), it will remain unclear whether the organization should invest in a

particular QI strategy, even if research shows demonstrable effects on QI. Future studies should incorporate cost, cost-effectiveness, or both in their research designs so as to be able to determine the value of the change for the organization and its patients.

Although a substantial proportion of the published studies report magnitude of change resulting from the QI intervention, the approach to assessing effect size varies. For example, there is considerable variation across studies in how magnitude is expressed (e.g., mean, median, correlation, regression coefficient), even across studies focused on similar types of QI interventions, which makes comparisons across these studies difficult. It is also the case that some studies report statistically significant results but show very small incremental improvements in quality while other studies report null findings (perhaps because they are underpowered) but show potentially large, practical QI changes. For example, Pandey and Cursio (2006) conducted a study to examine whether data feedback improved the quality of stroke care among 13 hospitals. They found measurable improvements in all seven outcome measures; however, because of the small sample of hospitals, only one outcome reached statistical significance.

Perhaps the most common concern expressed about the QI effectiveness studies is the extent to which the findings of these studies are generalizable to settings other than those in which the study was conducted. Much of this problem may be attributable to the customized effort that goes into many QI initiatives. For example, an important continuous QI principle is the ability to refine production processes on the basis of real-time data. Such efforts entail defining problems at a local level and developing solutions that are often unique to the situation and problems identified. This concern is exacerbated by the fact that many of the QI studies reviewed were done in single organizations, making the issue of generalizability more problematic. For example, a study by Serisier and Bowler (2007) measured the effect of an educational program targeting rapid and appropriate antibiotic administration on the management of community-acquired pneumonia in one hospital's emergency department. While their results showed a positive effect, the intervention was based on locally identified deficiencies in the process of community-acquired pneumonia care.

The lack of replication of QI studies also contributes to this concern. Until sufficiently large samples of organizations are included in QI effectiveness studies or until context is specifically considered as an integral influence on the appropriateness (or lack thereof) of a particular QI strategy, the issue of application of QI effectiveness findings will remain unresolved. It is important to note, however, that sample size considerations need to be balanced against the time and cost associated with recruitment and data collection across multiple sites versus single-site studies. For example, sample size requirements for larger studies that may offer greater potential for generalizability may be considerable, and quasi experiments may not be as affordable as true randomized trials involving single organizational samples and smaller numbers of participants.

Although single-organization samples are common in hospital QI studies, it is often the case that a number of subunits are employed in such studies. For example, a QI intervention might be introduced in a medical–surgical unit, and other such units in the hospital used as controls. However, this approach may raise concerns about the integrity of the research or potential ethical conflicts. For example, within a single organization, it may be difficult to limit the effects of contamination across subunits if staff has regular contact with each other. Similarly it may be difficult to introduce controls in such research if withholding a potentially valuable intervention would potentially deprive patients of needed assistance.

Recommendations for Future Research

Several findings of this literature review set the stage for discussion of the future of research about QI and quality outcomes. The first finding is that such a small number of published articles (185 of 7,842) from those yielded in the initial search met the criteria of being empirical, dealing with purposive organizational change aimed at improving quality, and having a dependent variable related to health care quality. The QI literature is clearly large and lacking empirical and theoretical coherence. This is surprising given the extensive discussion in the Institute of Medicine reports and in the systems approach advocated for reducing medical errors and improving quality of care (Joshi, Anderson, & Marwaha, 2002; Meyer & Massagli, 2001; Moray, 1994). Although the signal-to-noise ratio in our search was small, we do not recommend abandoning descriptive research or policy/management prescriptive literature, which no doubt has a place in the QI literature. However, if we researchers hope to improve our understanding of the effects of QI on quality outcomes, we need to be able to discern what is supported by empirical research from that which is simply prescribed as important. Doing so will require precise definitions, categorization, and operationalization of QI interventions by researchers.

A second important finding involves the variety in the ways QI is defined and operationalized across different studies. For example, a variable such as *teams* was conceptualized and measured in a number of ways, including team structure, team communication, team personnel, and team climate. This variability reflects the problems identified earlier and, perhaps more important, the absence of a solid conceptual framework that would help identify the key elements of a specific QI change, as well as the relevant outcome for that change. Theoretical and conceptual frameworks relevant to group relationships, organizational change, organizational learning, or adoption and implementation of innovations are only a few examples of those potentially relevant to QI research. Clearly the use of such frameworks would help greatly to organize a literature that is expanding in many disparate directions.

In terms of designing more useful studies, research should seek to address threats to external validity that may result from interactions between intervention characteristics and contextual factors in settings to which one might wish to generalize. This

involves considering the extent to which the study results can be generalized from the specific conditions in the study to various defined populations, other variations of the intervention, other outcomes, or other settings in context. In practical terms this might involve replications of a QI study that involve the use of variants of the intervention, different study contexts, or different experimental participants. For example, Marrie et al. (2000) examined whether the implementation of a critical pathway was associated with a reduction in community-acquired pneumonia in both community and teaching hospitals in Canada. As opposed to most RCTs, in which patients are randomized, this study randomized hospitals to the control and intervention groups, thereby allowing the researchers to control for confounding factors associated with hospital type. Likewise, a study by Daucourt, Saillour-Glenisson, Michel, Jutand, and Abouelfath (2003) examined the effectiveness of two different types of thyroid function guidelines. This study examined the effectiveness of these guidelines in small, midsize, and psychiatric hospitals, as well as across different types of units within these hospitals (e.g., general medicine, emergency room, and psychiatry).

Finally, few studies examined the interrelationships of different, supposedly interrelated QI changes to test the effects of systems of organizational factors on quality. Although many of the studies reviewed considered multiple QI changes, these changes were typically not evaluated from a systems theory perspective. The systems approach cited as pivotal to solving patient safety and quality problems by publications such as *To Err Is Human* (Institute of Medicine, 2000) and *Crossing the Quality Chasm* (Institute of Medicine, 2001) encourages health services researchers to consider greater application of systems-focused theory to questions of how organizational factors shape quality. A systems approach views outcomes such as quality and errors as the result of interconnected processes and parts that combine to address some common purpose (Institute of Medicine, 2001). Health services researchers can use these models to better understand how particular quality problems serve as examples of general organizational or work situations. This would help link seemingly disparate types of quality outcomes together, allow different quality outcomes to be studied simultaneously, and allow research results to be generalized more easily in the QI arena (Hoff, Jameson, Hannan, & Flink, 2004). To achieve more systems focus in QI effectiveness research, study designs would benefit from a multidisciplinary approach. Many QI studies have been designed and conducted by physician investigators with little or no input from economists, organizational behaviorists, or other related disciplines. The narrow focus on medical aspects of QI ignore the critical roles of organizational context, cost-effectiveness, and perhaps most important, the value added by the QI intervention to the patient or organization. Although the logistical considerations of multidisciplinary research are not trivial, the end product is likely to yield findings of greater relevance and utility than those conducted within a single disciplinary perspective.

Appendix

Reviewed Articles

Achterberg, W. P., Holtkamp, C. C., Kerkstra, A., Pot, A. M., Ooms, M. E., & Ribbe, M. W. (2001). Improvements in the quality of co-ordination of nursing care following implementation of the Resident Assessment Instrument in Dutch nursing homes. *Journal of Advanced Nursing*, *35*(2), 268-275.

Andrews, J. O., Tingen, M. S., Waller, J. L., & Harper, R. J. (2001). Provider feedback improves adherence with AHCPR smoking cessation guideline. *Preventive Medicine*, *33*(5), 415-421.

Ansari, F., Gray, K., Nathwani, D., Phillips, G., Ogston, S., Ramsay, C., et al. (2003). Outcomes of an intervention to improve hospital antibiotic prescribing: Interrupted time series with segmented regression analysis. *Journal of Antimicrobial Chemotherapy*, *52*(5), 842-848.

Armstrong, E. P. (2001). Clinical and economic outcomes of an ambulatory urinary tract infection disease management program. *American Journal of Managed Care*, *7*, 269-280.

Aubert, R. E., Herman, W. H., Waters, J., Moore, W., Sutton, D., Peterson, B. L., et al. (1998). Nurse case management to improve glycemic control in diabetic patients in a health maintenance organization: A randomized, controlled trial. *Annals of Internal Medicine*, *129*(8), 605-612.

Babcock, H. M., Zack, J. E., Garrison, T., Trovillion, E., Jones, M., Fraser, V. J., et al. (2004). An educational intervention to reduce ventilator-associated pneumonia in an integrated health system: A comparison of effects. *Chest*, *125*(6), 2224-2231.

Baker, R., Smith, J. F., & Lambert, P. C. (2003). Randomised controlled trial of the effectiveness of feedback in improving test ordering in general practice. *Scandinavian Journal of Primary Health Care*, *21*(4), 219-223.

Bates, D. W., Teich, J. M., Lee, J., Seger, D., Kuperman, G. J., Ma'Luf, N., et al. (1999). The impact of computerized physician order entry on medication error prevention. *Journal of the American Medical Informatics Association*, *6*(4), 313-321.

Berenholtz, S. M., Pronovost, P. J., Lipsett, P. A., Hobson, D., Earsing, K., Farley, J. E., et al. (2004). Eliminating catheter-related bloodstream infections in the intensive care unit. *Critical Care Medicine*, *32*(10), 2014-2020.

Berild, D., Ringertz, S. H., Aabyholm, G., Lelek, M., & Fosse, B. (2002). Impact of an antibiotic policy on antibiotic use in a paediatric department: Individual based follow-up shows that antibiotics were chosen according to diagnoses and bacterial findings. *International Journal of Antimicrobial Agents*, *20*(5), 333-338.

Berner, E. S., Baker, C. S., Funkhouser, E., Heudebert, G. R., Allison, J. J., Fargason, C. A., Jr., et al. (2003). Do local opinion leaders augment hospital quality improvement efforts? A randomized trial to promote adherence to unstable angina guidelines. *Medical Care*, *41*(3), 420-431.

Bexell, A., Lwando, E., von Hofsten, B., Tembo, S., Eriksson, B., & Diwan, V. K. (1996). Improving drug use through continuing education: A randomized controlled trial in Zambia. *Journal of Clinical Epidemiology*, *49*(3), 355-357.

Bickell, N. A., Zdeb, M. S., Applegate, M. S., Roohan, P. J., & Siu, A. L. (1996). Effect of external peer review on cesarean delivery rates: A statewide program. *Obstetrics & Gynecology*, *87*(Part 1), 664-667.

Bischoff, W. E., Reynolds, T. M., Sessler, C. N., Edmond, M. B., & Wenzel, R. P. (2000). Handwashing compliance by health care workers: The impact of introducing an accessible, alcohol-based hand antiseptic. *Archives of Internal Medicine*, *160*(7), 1017-1021.

Bours, G. J., Halfens, R. J., Candel, M. J., Grol, R. T., & Abu-Saad, H. H. (2004). A pressure ulcer audit and feedback project across multi-hospital settings in the Netherlands. *International Journal for Quality in Health Care*, *16*(3), 211-218.

Brummel-Smith, K., Joseph, A., & Boult, C. (2004). Managed primary care of nursing home residents. *Journal of the American Geriatrics Society*, *46*(9), 1152-1156.

(continued)

Appendix (continued)

Cabana, M. D., Slish, K. K., Evans, D., Mellins, R. B., Brown, R. W., Lin, X., et al. (2006). Impact of physician asthma care education on patient outcomes. *Pediatrics*, *117*(6), 2149-2157.

Caputo, R. P., Kosinski, R., Walford, G., Giambartolomei, A., Grant, W., Reger, M. J., et al. (2005). Effect of continuous quality improvement analysis on the delivery of primary percutaneous revascularization for acute myocardial infarction: A community hospital experience. *Catheterization and Cardiovascular Interventions*, *64*(4), 428-433; discussion 434-435.

Carling, P., Fung, T., Killion, A., Terrin, N., & Barza, M. (2003). Favorable impact of a multidisciplinary antibiotic management program conducted during 7 years. *Infection Control and Hospital Epidemiology*, *24*(9), 699-706.

Caron, A., & Neuhauser, D. (1999). The effect of public accountability on hospital performance: Trends in rates for cesarean sections and vaginal births after cesarean section in Cleveland, Ohio. *Quality Management in Health Care*, *7*(2), 1-10.

Cates, C. (1999). An evidence based approach to reducing antibiotic use in children with acute otitis media: Controlled before and after study. *British Medical Journal*, *318*(7185), 715-716.

Chalker, J., Ratanawijitrasin, S., Chuc, N. T. K., Petzold, M., & Tomson, G. (2005). Effectiveness of a multi-component intervention on dispensing practices at private pharmacies in Vietnam and Thailand: A randomized controlled trial. *Social Science & Medicine*, *60*(1), 131-141.

Charbonneau, A., Parker, V., Meterko, M., Rosen, A. K., Kader, B., Owen, R. R., et al. (2004). The relationship of system-level quality improvement with quality of depression care. *American Journal of Managed Care*, *10*(11 Pt. 2), 846-851.

Christakis, D. A., Zimmerman, F. J., Wright, J. A., Garrison, M. M., Rivara, F. P., & Davis, R. L. (2001). A randomized controlled trial of point-of-care evidence to improve the antibiotic prescribing practices for otitis media in children. *Pediatrics*, *107*(2), e15-e19.

Chu, L. A., Bratzler, D. W., Lewis, R. J., Murray, C., Moore, L., Shook, C., et al. (2003). Improving the quality of care for patients with pneumonia in very small hospitals. *Archives of Internal Medicine*, *163*(3), 326-332.

Chuc, N. T., Larsson, M., Do, N. T., Diwan, V. K., Tomson, G. B., & Falkenberg, T. (2002). Improving private pharmacy practice: A multi-intervention experiment in Hanoi, Vietnam. *Journal of Clinical Epidemiology*, *55*(11), 1148-1155.

Climo, M. W., Israel, D. S., Wong, E. S., Williams, D., Coudron, P., & Markowitz, S. M. (1998). Hospital-wide restriction of clindamycin: Effect on the incidence of clostridium difficile-associated diarrhea and cost. *Annals of Internal Medicine*, *128*(12 Part 1), 989-995.

Coopersmith, C. M., Zack, J. E., Ward, M. R., Sona, C. S., Schallom, M. E., Everett, S. J., et al. (2004). The impact of bedside behavior on catheter-related bacteremia in the intensive care unit. *Archives of Surgery*, *139*, 131-136.

Couch, C., Sheffield, P., Gerthoffer, T., Ries, A., & Hollander, P. (2003). Clinical outcomes in patients with type 2 diabetes managed by a diabetes resource nurse in a primary care practice. *Baylor University Medical Center Proceedings*, *16*(3), 336-340.

Daucourt, V., Saillour-Glenisson, F., Michel, P., Jutand, M. A., & Abouelfath, A. (2003). A multicenter cluster randomized controlled trial of strategies to improve thyroid function testing. *Medical Care*, *41*(3), 432-441.

Dean, N. C., Silver, M. P., Bateman, K. A., James, B., Hadlock, C. J., & Hale, D. (2001). Decreased mortality after implementation of a treatment guideline for community-acquired pneumonia. *American Journal of Medicine*, *110*(6), 451-457.

Debehnke, D., & Decker, M. C. (2002). The effects of a physician–nurse patient care team on patient satisfaction in an academic ED. *American Journal of Emergency Medicine*, *20*(4), 267-270.

(continued)

Appendix (continued)

Dedrick, R. E., Sinkowitz-Cochran, R. L., Cunningham, C., Muder, R. R., Perreiah, P., Cardo, D. M., et al. (2007). Hand hygiene practices after brief encounters with patients: An important opportunity for prevention. *Infection Control and Hospital Epidemiology*, *28*(3), 341-345.

Del Beccaro, M. A., Jeffries, H. E., Eisenberg, M. A., & Harry, E. D. (2006). Computerized provider order entry implementation: No association with increased mortality rates in an intensive care unit. *Pediatrics*, *118*(1), 290-295.

de Rond, M., de Wit, R., & van Dam, F. (2001). The implementation of a pain monitoring programme for nurses in daily clinical practice: Results of a follow-up study in five hospitals. *Journal of Advanced Nursing*, *35*(4), 590-598.

Dietrich, A. J., O'Connor, G. T., Keller, A., Carney, P. A., Levy, D., & Whaley, F. S. (1992). Cancer: Improving early detection and prevention: A community practice randomised trial. *British Medical Journal*, *304*(6828), 687-691.

Dietrich, A. J., Oxman, T. E., Williams, J. W., Jr., Schulberg, H. C., Bruce, M. L., Lee, P. W., et al. (2004). Re-engineering systems for the treatment of depression in primary care: Cluster randomised controlled trial. *British Medical Journal*, *329*(7466), 602-604.

Diwan, V. K., Wahlstrom, R., Tomson, G., Beermann, B., Sterky, G., & Eriksson, B. (1995). Effects of "group detailing" on the prescribing of lipid-lowering drugs: A randomized controlled trial in Swedish primary care. *Journal of Clinical Epidemiology*, *48*(5), 705-711.

Dolan, N. C., Ng, J. S., Martin, G. J., Robinson, J. K., & Rademaker, A. W. (1997). Effectiveness of a skin cancer control educational intervention for internal medicine housestaff and attending physicians. *Journal of General Internal Medicine*, *12*(9), 531-536.

Dranitsaris, G., Spizzirri, D., Pitre, M., & McGeer, A. (2001). A randomized trial to measure the optimal role of the pharmacist in promoting evidence-based antibiotic use in acute care hospitals. *International Journal of Technology Assessment in Health Care*, *17*(2), 171-180.

Eccles, M., McColl, E., Steen, N., Rousseau, N., Grimshaw, J., Parkin, D., et al. (2002). Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: Cluster randomised controlled trial. *British Medical Journal*, *325*(7370), 941-944.

Eccles, M., Steen, N., Grimshaw, J., Thomas, L., McNamee, P., Soutter, J., et al. (2001). Effect of audit and feedback, and reminder messages on primary-care radiology referrals: A randomised trial. *Lancet*, *357*(9266), 1406-1409.

Eggimann, P., Harbarth, S., Constantin, M. N., Touveneau, S., Chevrolet, J. C., & Pittet, D. (2000). Impact of a prevention strategy targeted at vascular-access care on incidence of infections acquired in intensive care. *Lancet*, *355*(9218), 1864-1868.

Eldridge, N. E., Woods, S. S., Bonello, R. S., Clutter, K., Ellingson, L., Harris, M. A., et al. (2006). Using the six sigma process to implement the Centers for Disease Control and Prevention Guideline for Hand Hygiene in 4 intensive care units. *Journal of General Internal Medicine*, *21*(Suppl. 2), S35-42.

Fairbrother, G., Hanson, K. L., Friedman, S., & Butts, G. C. (1999). The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates. *American Journal of Public Health*, *89*(2), 171-175.

Farris, K. B., Kirking, D. M., Shimp, L. A., & Opdycke, R. A. (1996). Design and results of a group counter-detailing DUR educational program. *Pharmaceutical Research*, *13*(10), 1445-1452.

Ferguson, T. B., Jr., Peterson, E. D., Coombs, L. P., Eiken, M. C., Carey, M. L., Grover, F. L., et al. (2003). Use of continuous quality improvement to increase use of process measures in patients undergoing coronary artery bypass graft surgery: A randomized controlled trial. *Journal of the American Medical Association*, *290*(1), 49-56.

Feucht, C. L., & Rice, L. B. (2003). An interventional program to improve antibiotic use. *Annals of Pharmacotherapy*, *37*(5), 646-651.

(continued)

Appendix (continued)

Finkelstein, J. A., Davis, R. L., Dowell, S. F., Metlay, J. P., Soumerai, S. B., Rifas-Shiman, S. L., et al. (2001). Reducing antibiotic use in children: A randomized trial in 12 practices. *Pediatrics*, *108*(1), 1-7.

Finnema, E., de Lange, J., Droes, R. M., Ribbe, M., & van Tilburg, W. (2001). The quality of nursing home care: Do the opinions of family members change after implementation of emotion-oriented care? *Journal of Advanced Nursing*, *35*(5), 728-740.

Flottorp, S., Oxman, A. D., Havelsrud, K., Treweek, S., & Herrin, J. (2002). Cluster randomised controlled trial of tailored interventions to improve the management of urinary tract infections in women and sore throat. *British Medical Journal*, *325*(7360), 367-369.

Fogarty, G. J., & Mckeon, C. M. (2006). Patient safety during medication administration: The influence of organizational and individual variables on unsafe work practices and medication errors. *Ergonomics*, *49*(5), 444-456.

Frances, C. D., Alperin, P., Adler, J. S., & Grady, D. (2001). Does a fixed physician reminder system improve the care of patients with coronary artery disease? A randomized controlled trial. *Western Journal of Medicine*, *175*(3), 165-166.

Gall, M. J., Harmer, J. E., & Wanstall, H. J. (2001). Prescribing of oral nutritional supplements in primary care: Can guidelines supported by education improve prescribing practice? *Clinical Nutrition*, *20*(6), 511-515.

Gastmeier, P., Brauer, H., Forster, D., Dietz, E., Daschner, F., & Ruden, H. (2002). A quality management project in 8 selected hospitals to reduce nosocomial infections: A prospective, controlled study. *Infection Control and Hospital Epidemiology*, *23*(2), 91-97.

Gentry, C. A., Greenfield, R. A., Slater, L. N., Wack, M., & Huycke, M. M. (2000). Outcomes of an antimicrobial control program in a teaching hospital. *American Journal of Health System Pharmacy*, *57*(3), 268-274.

Gill, P., Ryan, J., Morgan, O., & Williams, A. (2000). Team nursing and ITU: A good combination? *Intensive and Critical Care Nursing*, *16*(4), 243-255.

Goff, D. C., Jr., Gu, L., Cantley, L. K., Parker, D. G., & Cohen, S. J. (2002). Enhancing the quality of care for patients with coronary heart disease: The design and baseline results of the Hastening the Effective Application of Research Through Technology (HEART) trial. *American Journal of Managed Care*, *8*, 1069-1078.

Goldberg, D., Schiff, G. D., McNutt, R., Furumoto-Dawson, A., Hammerman, M., & Hoffman, A. (2004). Mailings timed to patients' appointments: A controlled trial of fecal occult blood test cards. *American Journal of Preventive Medicine*, *26*(5), 431-435.

Goldspiel, B. R., DeChristoforo, R., & Daniels, C. E. (2000). A continuous-improvement approach for reducing the number of chemotherapy-related medication errors. *American Journal of Health System Pharmacy*, *57*(Suppl. 4), S4-9.

Gonzales, R., Steiner, J. F., Lum, A., & Barrett, P. H., Jr. (1999). Decreasing antibiotic use in ambulatory practice: Impact of a multidimensional intervention on the treatment of uncomplicated acute bronchitis in adults. *Journal of the American Medical Association*, *281*(16), 1512-1519.

Goode, C. J., Tanaka, D. J., Krugman, M., O'Connor, P. A., Bailey, C., Deutchman, M., et al. (2000). Outcomes from use of an evidence-based practice guideline. *Nursing Economics*, *18*(4), 202-207.

Greene, J. P., Rogers, V. W. M. D., & Yedidia, M. J. P. (2007). The impact of implementing a chronic care residency training initiative on asthma outcomes. *Academic Medicine*, *82*(2), 161-167.

Greineder, D. K., Loane, K. C., & Parks, P. (1999). A randomized controlled trial of a pediatric asthma outreach program. *Journal of Allergy and Clinical Immunology*, *103*(3 Pt. 1), 436-440.

Gross, R., Morgan, A. S., Kinky, D. E., Weiner, M., Gibson, G. A., & Fishman, N. O. (2001). Impact of a hospital-based antimicrobial management program on clinical and economic outcomes. *Clinical Infectious Diseases*, *33*(3), 289-295.

(continued)

Appendix (continued)

Grossbart, S. R. (2006). What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. *Medical Care Research and Review*, *63*(1 Suppl.), S29-S48.

Gums, J. G., Yancey, R. W., Jr., Hamilton, C. A., & Kubilis, P. S. (1999). A randomized, prospective study measuring outcomes after antibiotic therapy intervention by a multidisciplinary consult team. *Pharmacotherapy*, *19*(12), 1369-1377.

Gutierrez, G., Guiscafre, H., Bronfman, M. S., Walsh, J., Martinez, H., & Munoz, O. (1994). Changing physician prescribing patterns: Evaluation of an educational strategy for acute diarrhea in Mexico City. *Medical Care*, *32*(5), 436-446.

Guttman, A., Afilalo, M., Guttman, R., Colacone, A., Robitaille, C., Lang, E., et al. (2004). An emergency department-based nurse discharge coordinator for elder patients: Does it make a difference? *Academic Emergency Medicine*, *11*(12), 1318-1327.

Gyssens, I. C., Geerligs, I. E., Nannini-Bergman, M. G., Knape, J. T., Hekster, Y. A., & van der Meer, J. W. (1996). Optimizing the timing of antimicrobial prophylaxis in surgery: An intervention study. *Journal of Antimicrobial Chemotherapy*, *38*(2), 301-308.

Han, Y. Y., Carcillo, J. A., Venkataraman, S. T., Clark, R. S., Watson, R. S., Nguyen, T. C., et al. (2005). Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. *Pediatrics*, *116*(6), 1506-1512.

Hannan, E. L., Sarrazin, M. S., Doran, D. R., & Rosenthal, G. E. (2003). Provider profiling and quality improvement efforts in coronary artery bypass graft surgery: The effect on short-term mortality among Medicare beneficiaries. *Medical Care*, *41*(10), 1164-1172.

Harbarth, S., Pittet, D., Grady, L., Zawacki, A., Potter-Bynoe, G., Samore, M. H., et al. (2002). Interventional study to evaluate the impact of an alcohol-based hand gel in improving hand hygiene compliance. *Pediatric Infectious Disease Journal*, *21*(6), 489-495.

Harris, R. H., MacKenzie, T. D., Leeman-Castillo, B., Corbett, K. K., Batal, H. A., Maselli, J. H., et al. (2003). Optimizing antibiotic prescribing for acute respiratory tract infections in an urban urgent care clinic. *Journal of General Internal Medicine*, *18*(5), 326-334.

Harrison, R. V., Janz, N. K., Wolfe, R. A., Tedeschi, P. J., Chernew, M., Stross, J. K., et al. (2003). Personalized targeted mailing increases mammography among long-term noncompliant Medicare beneficiaries: A randomized trial. *Medical Care*, *41*(3), 375-385.

Haycock, C., Laser, C., Keuth, J., Montefour, K., Wilson, M., Austin, K., et al. (2005). Implementing evidence-based practice findings to decrease postoperative sternal wound infections following open heart surgery. *Journal of Cardiovascular Nursing*, *20*(5), 299-305.

Hayes, R. P., Baker, D. W., Luthi, J. C., Baggett, R. L., McClellan, W., Fitzgerald, D., et al. (2002). The effect of external feedback on the management of Medicare inpatients with congestive heart failure. *American Journal of Medical Quality*, *17*(6), 225-235.

Hennessy, T. W., Petersen, K. M., Bruden, D., Parkinson, A. J., Hurlburt, D., Getty, M., et al. (2002). Changes in antibiotic prescribing practices and carriage of penicillin-resistant Streptococcus pneumoniae: A controlled intervention trial in rural Alaska. *Clinical Infectious Diseases*, *34*, 1543-1550.

Hesselink, A. E., Penninx, B. W., van der Windt, D. A., van Duin, B. J., de Vries, P., Twisk, J. W., et al. (2004). Effectiveness of an education programme by a general practice assistant for asthma and COPD patients: Results from a randomised controlled trial. *Patient Education and Counseling*, *55*, 121-128.

Higuera, F., Rosenthal, V. D., Duarte, P., Ruiz, J., Franco, G., & Safdar, N. (2005). The effect of process control on the incidence of central venous catheter-associated bloodstream infections and mortality in intensive care units in Mexico. *Critical Care Medicine*, *33*(9), 2022-2027.

Hillman, A. L., Ripley, K., Goldfarb, N., Nuamah, I., Weiner, J., & Lusk, E. (1998). Physician financial incentives and feedback: Failure to increase cancer screening in Medicaid managed care. *American Journal of Public Health*, *88*(11), 1699-1701.

(continued)

Appendix (continued)

Horbar, J. D., Carpenter, J. H., Buzas, J., Soll, R. F., Suresh, G., Bracken, M. B., et al. (2004). Collaborative quality improvement to promote evidence based surfactant for preterm infants: A cluster randomised trial. *British Medical Journal*, *329*(7473), 1004-1007.

Hughes, M. R., Smith, C. D., Tecklenburg, F. W., Habib, D. M., Hulsey, T. C., & Ebeling, M. (2001). Effects of a weaning protocol on ventilated pediatric intensive care unit (PICU) patients. *Topics in Health Information Management*, *22*(2), 35-43.

Irvine Doran, D. M., Baker, G. R., Murray, M., Bohnen, J., Zahn, C., Sidani, S., et al. (2002). Achieving clinical improvement: An interdisciplinary intervention. *Health Care Management Review*, *27*(4), 42-56.

Jolly, K., Bradley, F., Sharp, S., Smith, H., Thompson, S., Kinmonth, A. L., et al. (1999). Randomised controlled trial of follow up care in general practice of patients with myocardial infarction and angina: Final results of the Southampton Heart Integrated Care Project (SHIP). *British Medical Journal*, *318*(7185), 706-711.

Joseph, A. M., Arikian, N. J. P., An, L. C., Nugent, S., Sloan, R. J., Pieper, C. F. D., et al. (2004). Results of a randomized controlled trial of intervention to implement smoking guidelines in Veterans Affairs Medical Centers: Increased use of medications without cessation benefit. *Medical Care*, *42*(11), 1100-1110.

Khan, R., & Cheesbrough, J. (2003). Impact of changes in antibiotic policy on clostridium difficile-associated diarrhoea (CDAD) over a five-year period in a district general hospital. *Journal of Hospital Infection*, *54*(2), 104-108.

Kiefe, C. I., Allison, J. J., Williams, O. D., Person, S. D., Weaver, M. T., & Weissman, N. W. (2001). Improving quality improvement using achievable benchmarks for physician feedback: A randomized controlled trial. *Journal of the American Medical Association*, *285*(22), 2871-2879.

Kimberlin, C. L., Berardo, D. H., Pendergast, J. F., & McKenzie, L. C. (1993). Effects of an education program for community pharmacists on detecting drug-related problems in elderly patients. *Medical Care*, *31*(5), 451-468.

Kinmonth, A. L., Woodcock, A., Griffin, S., Spiegal, N., & Campbell, M. J. (1998). Randomised controlled trial of patient centred care of diabetes in general practice: Impact on current wellbeing and future disease risk. *British Medical Journal*, *317*(7167), 1202-1208.

Kinsinger, L. S., Harris, R., Qaqish, B., Strecher, V., & Kaluzny, A. (1998). Using an office system intervention to increase breast cancer screening. *Journal of General Internal Medicine*, *13*(8), 507-514.

Kiwanuka, A. I., & Moore, W. M. (1993). Influence of audit and feedback on use of caesarean section in a geographically-defined population. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *50*(1), 59-64.

Landman, D., Chockalingam, M., & Quale, J. M. (1999). Reduction in the incidence of methicillin-resistant Staphylococcus aureus and ceftazidime-resistant Klebsiella pneumoniae following changes in a hospital antibiotic formulary. *Clinical Infectious Diseases*, *28*(5), 1062-1066.

Lankford, M. G., Zembower, T. R., Trick, W. E., Hacek, D. M., Noskin, G. A., & Peterson, L. R. (2003). Influence of role models and hospital design on hand hygiene of healthcare workers. *Emerging Infectious Diseases*, *9*(2), 217-223.

Lautenbach, E., LaRosa, L. A., Marr, A. M., Nachamkin, I., Bilker, W. B., & Fishman, N. O. (2003). Changes in the prevalence of vancomycin-resistant enterococci in response to antimicrobial formulary interventions: Impact of progressive restrictions on use of vancomycin and third-generation cephalosporins. *Clinical Infectious Diseases*, *36*(4), 440-446.

Legorreta, A. P., Peters, A. L., Ossorio, C., Lopez, R. J., Jatulis, D., & Davidson, M. B. (1996). Effect of a comprehensive nurse-managed diabetes program: An HMO prospective study. *American Journal of Managed Care*, *2*, 1024-1030.

Lemelin, J., Hogg, W., & Baskerville, N. (2001). Evidence to action: A tailored multifaceted approach to changing family physician practice patterns and improving preventive care. *Canadian Medical Association Journal*, *164*(6), 757-763.

(continued)

Appendix (continued)

Leverstein-van Hall, M., Fluit, A., Blok, H., Box, A., Peters, E., Weersink, A., et al. (2001). Control of nosocomial multiresistant enterobacteriaceae using a temporary restrictive antibiotic agent policy. *European Journal of Clinical Microbiology & Infectious Diseases*, *20*(11), 785-791.

Levin-Scherz, J., DeVita, N., & Timbie, J. (2006). Impact of pay-for-performance contracts and network registry on diabetes and asthma HEDIS measures in an integrated delivery network. *Medical Care Research and Review*, *63*(1 Suppl.), S14-S28.

Lindenauer, P. K., Remus, D., Roman, S., Rothberg, M. B., Benjamin, E. M., Ma, A., et al. (2007). Public reporting and pay for performance in hospital quality improvement. *New England Journal of Medicine*, *356*(5), 486-496.

Little, P., Gould, C., Williamson, I., Moore, M., Warner, G., & Dunleavey, J. (2001). Pragmatic randomised controlled trial of two prescribing strategies for childhood acute otitis media. *British Medical Journal*, *322*(7282), 336-342.

Lobo, R. D., Levin, A. S., Gomes, L. M., Cursino, R., Park, M., Figueiredo, V. B., et al. (2005). Impact of an educational program and policy changes on decreasing catheter-associated bloodstream infections in a medical intensive care unit in Brazil. *American Journal of Infection Control*, *33*(2), 83-87.

Lowensteyn, I., Joseph, L., Levinton, C., Abrahamowicz, M., Steinert, Y., & Grover, S. (1998). Can computerized risk profiles help patients improve their coronary risk? The results of the Coronary Health Assessment Study (CHAS). *Preventive Medicine*, *27*(5 Pt 1), 730-737.

Lozano, P., Finkelstein, J. A., Carey, V. J., Wagner, E. H., Inui, T. S., Fuhlbrigge, A. L., et al. (2004). A multisite randomized trial of the effects of physician education and organizational change in chronic-asthma care: Health outcomes of the Pediatric Asthma Care Patient Outcomes Research Team II Study. *Archives of Pediatric Adolescent Medicine*, *158*(9), 875-883.

Lundborg, C. S., Wahlstrom, R., Oke, T., Tomson, G., & Diwan, V. K. (1999). Influencing prescribing for urinary tract infection and asthma in primary care in Sweden: A randomized controlled trial of an interactive educational intervention. *Journal of Clinical Epidemiology*, *52*(8), 801-812.

Lynn, J., West, J., Hausmann, S., Gifford, D., Nelson, R., McGann, P., et al. (2007). Collaborative clinical quality improvement for pressure ulcers in nursing homes. *Journal of the American Geriatric Society*, *55*(10), 1663-1669.

Macfarlane, J., Holmes, W., Gard, P., Thornhill, D., Macfarlane, R., Hubbard, R., et al. (2002). Reducing antibiotic use for acute bronchitis in primary care: Blinded, randomised controlled trial of patient information leaflet. *British Medical Journal*, *324*(7329), 91-93.

Main, E. K. (1999). Reducing cesarean birth rates with data-driven quality improvement activities. *Pediatrics*, *103*(1), 374-383.

Mainous, A. G., III, Hueston, W. J., Love, M. M., Evans, M. E., & Finger, R. (2000). An evaluation of statewide strategies to reduce antibiotic overuse. *Family Medicine*, *32*(1), 22-29.

Marciniak, T. A., Ellerbeck, E. F., Radford, M. J., Kresowik, T. F., Gold, J. A., Krumholz, H. M., et al. (1998). Improving the quality of care for Medicare patients with acute myocardial infarction: Results from the Cooperative Cardiovascular Project. *Journal of the American Medical Association*, *279*(17), 1351-1357.

Margolis, P. A., Stevens, R., Bordley, W. C., Stuart, J., Harlan, C., Keyes-Elstein, L., et al. (2001). From concept to application: The impact of a community-wide intervention to improve the delivery of preventive services to children. *Pediatrics*, *108*(3), e42-e48.

Marrie, T. J., Lau, C. Y., Wheeler, S. L., Wong, C. J., Vandervoort, M. K., Feagan, B. G., et al. (2000). A controlled trial of a critical pathway for treatment of community-acquired pneumonia. *Journal of the American Medical Association*, *283*(6), 749-755.

(continued)

Appendix (continued)

Martinez, M. J., Freire, A., Castro, I., Inaraja, M. T., Ortega, A., Del Campo, V., et al. (2000). Clinical and economic impact of a pharmacist-intervention to promote sequential intravenous to oral clindamycin conversion. *Pharmacy World & Science*, *22*(2), 53-58.

Mayer, T. A., Cates, R. J., Mastorovich, M. J., & Royalty, D. L. (1998). Emergency department patient satisfaction: Customer service training improves patient satisfaction and ratings of physician and nurse skill. *Journal of Healthcare Management*, *43*(5), 427-440.

Mazor, S. S., Hampers, L. C., Chande, V. T., & Krug, S. E. (2002). Teaching Spanish to pediatric emergency physicians: Effects on patient satisfaction. *Archives of Pediatric Adolescent Medicine*, *156*(7), 693-695.

McCartney, P., Macdowall, W., & Thorogood, M. (1997). A randomised controlled trial of feedback to general practitioners of their prophylactic aspirin prescribing. *British Medical Journal*, *315*(7099), 35-36.

McIsaac, W. J., Goel, V., To, T., Permaul, J. A., & Low, D. E. (2002). Effect on antibiotic prescribing of repeated clinical prompts to use a sore throat score: Lessons from a failed community intervention study. *Journal of Family Practice*, *51*(4), 339-345.

McNulty, C., Logan, M., Donald, I., Ennis, D., Taylor, D., Baldwin, R., et al. (1997). Successful control of Clostridium difficile infection in an elderly care unit through use of a restrictive antibiotic policy. *Journal of Antimicrobial Chemotherapy*, *40*(5), 707-711.

Mehta, R., Das, S., Tsai, T. T., Nolan, E., Kearly, G., & Eagle, K. A. (2001). Quality improvement initiative and its impact on the management of patients with acute myocardial infarction. *Archives of Internal Medicine*, *160*(20), 3057-3062.

Melander, E., Bjorgell, A., Bjorgell, P., Ovhed, I., & Molstad, S. (1999). Medical audit changes physicians prescribing of antibiotics for respiratory tract infections. *Scandinavian Journal of Primary Health Care*, *17*(3), 180-184.

Merlani, P., Garnerin, P., Diby, M., Ferring, M., & Ricou, B. (2001). Quality improvement report: Linking guideline to regular feedback to increase appropriate requests for clinical tests: Blood gas analysis in intensive care. *British Medical Journal*, *323*(7313), 620-624.

Moher, M., Yudkin, P., Wright, L., Turner, R., Fuller, A., Schofield, T., et al. (2001). Cluster randomised controlled trial to compare three methods of promoting secondary prevention of coronary heart disease in primary care. *British Medical Journal*, *322*(7298), 1338-1347.

Monteleoni, C., & Clark, E. (2004). Using rapid-cycle quality improvement methodology to reduce feeding tubes in patients with advanced dementia: Before and after study. *British Medical Journal*, *329*, 491-494.

Mullen, B. A., & Watts Kelley, P. A. (2006). Diabetes nurse case management: An effective tool. *Journal of the American Academy of Nurse Practitioners*, *18*(1), 22-30.

Munck, A. P., Gahrn-Hansen, B., Sogaard, P., & Sogaard, J. (1999). Long-lasting improvement in general practitioners' prescribing of antibiotics by means of medical audit. *Scandinavian Journal of Primary Health Care*, *17*(3), 185-190.

Naughton, B. J., Mylotte, J. M., Ramadan, F., Karuza, J., & Priore, R. L. (2001). Antibiotic use, hospital admissions, and mortality before and after implementing guidelines for nursing home-acquired pneumonia. *Journal of the American Geriatrics Society*, *49*(8), 1020-1024.

O'Connell, D. L., Henry, D., & Tomlins, R. (1999). Randomised controlled trial of effect of feedback on general practitioners' prescribing in Australia. *British Medical Journal*, *318*(7182), 507-511.

Pandey, D., & Cursio, J. (2006). Data feedback for quality improvement of stroke care CAPTURE stroke experience. *American Journal of Preventive Medicine*, *31*(6), S224-S229.

Perz, J. F., Craig, A. S., Coffey, C. S., Jorgensen, D. M., Mitchel, E., Hall, S., et al. (2002). Changes in antibiotic prescribing for children after a community-wide campaign. *Journal of the American Medical Association*, *287*(23), 3103-3109.

(continued)

Appendix (continued)

Peterson, E. D., Roe, M. T., Mulgund, J., DeLong, E. R., Lytle, B. L., Brindis, R. G., et al. (2006). Association between hospital process performance and outcomes among patients with acute coronary syndromes. *Journal of the American Medical Association*, *295*(16), 1912-1920.

Pimlott, N. J. G., Hux, J. E., Wilson, L. M., Kahan, M., Li, C., & Rosser, W. W. (2003). Educating physicians to reduce benzodiazepine use by elderly patients: A randomized controlled trial. *Canadian Medical Association Journal*, *168*(7), 835-839.

Poma, P. A. (1998). Effect of departmental policies on cesarean delivery rates: A community hospital experience. *Obstetrics & Gynecology*, *91*(6), 1013-1018.

Rakowski, W., Ehrich, B., Goldstein, M. G., Rimer, B. K., Pearlman, D. N., Clark, M. A., et al. (1998). Increasing mammography among women aged 40-74 by use of a stage-matched, tailored intervention. *Preventive Medicine*, *27*(5 Pt. 1), 748-756.

Ramunno, L. D., Dodds, T. A., & Traven, N. D. (1998). Cooperative Cardiovascular Project (CCP) quality improvement in Maine, New Hampshire, and Vermont. *Evaluation & the Health Professions* , *21*(4), 442-460.

Rask, K., Parmelee, P. A., Taylor, J. A., Green, D., Brown, H., Hawley, J., et al. (2007). Implementation and evaluation of a nursing home fall management program. *Journal of the American Geriatrics Society*, *55*(3), 342-349.

Reddy, J. C., Katz, P. P., Goldman, L., & Wachter, R. M. (2001). A pneumonia practice guideline and a hospitalist-based reorganization lead to equivalent efficiency gains. *American Journal of Managed Care*, *7*(12), 1142-1148.

Rhew, D. C., Glassman, P. A., & Goetz, M. B. (1999). Improving pneumococcal vaccine rates: Nurse protocols versus clinical reminders. *Journal of General Internal Medicine*, *14*(6), 351-356.

Richards, M. J., Robertson, M. B., Dartnell, J. G., Duarte, M. M., Jones, N. R., Kerr, D. A., et al. (2003). Impact of a web-based antimicrobial approval system on broad-spectrum cephalosporin use at a teaching hospital. *Medical Journal of Australia*, *178*, 386-390.

Robson, M. S., Scudamore, I. W., & Walsh, S. M. (1996). Using the medical audit cycle to reduce cesarean section rates. *American Journal of Obstetrics and Gynecology*, *174*(1), 199-205.

Rokstad, K., Straand, J., & Fugelli, P. (1995). Can drug treatment be improved by feedback on prescribing profiles combined with therapeutic recommendations? A prospective, controlled trial in general practice. *Journal of Clinical Epidemiology*, *48*(8), 1061-1068.

Rollman, B. L., Hanusa, B. H., Lowe, H. J., Gilbert, T., Kapoor, W. N., & Schulberg, H. C. (2002). A randomized trial using computerized decision support to improve treatment of major depression in primary care. *Journal of General Internal Medicine*, *17*(7), 493-503.

Rosenthal, V. D., Guzman, S., Pezzotto, S. M., & Crnich, C. J. (2003). Effect of an infection control program using education and performance feedback on rates of intravascular device-associated bloodstream infections in intensive care units in Argentina. *American Journal of Infection Control*, *31*(7), 405-409.

Rosenthal, V. D., Guzman, S., & Safdar, N. (2004). Effect of education and performance feedback on rates of catheter-associated urinary tract infection in intensive care units in Argentina. *Infection Control and Hospital Epidemiology*, *25*(1), 47-50.

Rosenthal, V. D., Guzman, S., & Safdar, N. (2005). Reduction in nosocomial infection with improved hand hygiene in intensive care units of a tertiary care hospital in Argentina. *Infection Control and Hospital Epidemiology*, *33*(7), 392-397.

Ross-Degnan, D., Soumerai, S. B., Goel, P. K., Bates, J., Makhulo, J., Dondi, N., et al. (1996). The impact of face-to-face educational outreach on diarrhoea treatment in pharmacies. *Health Policy and Planning*, *11*(3), 308-318.

(continued)

Appendix (continued)

Saizy-Callaert, S., Causse, R., Furhman, C., Le Paih, M. F., Thebault, A., & Chouaid, C. (2003). Impact of a multidisciplinary approach to the control of antibiotic prescription in a general hospital. *Journal of Hospital Infection*, *53*(3), 177-182.

Salahuddin, N., Zafar, A., Sukhyani, L., Rahim, S., Noor, M. F., Hussain, K., et al. (2004). Reducing ventilator-associated pneumonia rates through a staff education programme. *Journal of Hospital Infection*, *57*(3), 223-227.

Santoso, B. (1996). Small group intervention vs formal seminar for improving appropriate drug use. *Social Science & Medicine*, *42*(8), 1163-1168.

Schectman, J. M., Schroth, W. S., Verme, D., & Voss, J. D. (2003). Randomized controlled trial of education and feedback for implementation of guidelines for acute low back pain. *Journal of General Internal Medicine*, *18*(10), 773-780.

Schelenz, S., Tucker, D., Georgeu, C., Daly, S., Hill, M., Roxburgh, J., et al. (2005). Significant reduction of endemic MRSA acquisition and infection in cardiothoracic patients by means of an enhanced targeted infection control programme. *Journal of Hospital Infection*, *60*(2), 104-110.

Schentag, J. J., Ballow, C. H., Fritz, A. L., Paladino, J. A., Williams, J. D., Cumbo, T. J., et al. (1993). Changes in antimicrobial agent usage resulting from interactions among clinical pharmacy, the infectious disease division, and the microbiology laboratory. *Diagnostic Microbiology & Infectious Disease*, *16*(3), 255-264.

Serisier, D. J., & Bowler, S. D. (2007). Effect of a simple educational intervention on the hospital management of community-acquired pneumonia. *Respirology*, *12*, 389-393.

Shannon, K. C., Sinacore, J. M., Bennett, S. G., Joshi, A. M., Sherin, K. M., & Deitrich, A. (2001). Improving delivery of preventive health care with the Comprehensive Annotated Reminder Tool (CART). *Journal of Family Practice*, *50*(9), 767-771.

Sherertz, R. J., Ely, E. W., Westbrook, D. M., Gledhill, K. S., Streed, S. A., Kiger, B., et al. (2000). Education of physicians-in-training can decrease the risk for vascular catheter infection. *Annals of Internal Medicine*, *132*(8), 641-648.

Shesser, R., Smith, M., Adams, S., Walls, R., & Paxton, M. (1986). The effectiveness of an organized emergency department follow-up system. *Annals of Emergency Medicine*, *15*(8), 911-915.

Shojania, K. G., Yokoe, D., Platt, R., Fiskio, J., Ma'luf, N., & Bates, D. W. (1998). Reducing vancomycin use utilizing a computer guideline: Results of a randomized controlled trial. *Journal of the American Medical Informatics Association*, *5*(6), 554-562.

Simon, G. E., VonKorff, M., Rutter, C., & Wagner, E. (2000). Randomised trial of monitoring, feedback, and management of care by telephone to improve treatment of depression in primary care. *British Medical Journal*, *320*(7234), 550-554.

Smyrnios, N. A., Connolly, A., Wilson, M. M., Curley, F. J., French, C. T., Heard, S. O., et al. (2002). Effects of a multifaceted, multidisciplinary, hospital-wide quality improvement program on weaning from mechanical ventilation. *Critical Care Medicine*, *30*(6), 1224-1230.

Solomon, D. H., Van Houten, L., Glynn, R. J., Baden, L., Curtis, K., Schrager, H., et al. (2001). Academic detailing to improve use of broad-spectrum antibiotics at an academic medical center. *Archives of Internal Medicine*, *161*(15), 1897-1902.

Sondergaard, J., Andersen, M., Stovring, H., & Kragstrup, J. (2003). Mailed prescriber feedback in addition to a clinical guideline has no impact: A randomised, controlled trial. *Scandinavian Journal of Primary Health Care*, *21*(1), 47-51.

Soumerai, S. B., McLaughlin, T. J., Gurwitz, J. H., Guadagnoli, E., Hauptman, P. J., Borbas, C., et al. (1998). Effect of local medical opinion leaders on quality of care for acute myocardial infarction: A randomized controlled trial. *Journal of the American Medical Association*, *279*(17), 1358-1363.

(continued)

Appendix (continued)

Spinewine, A., Swine, C., Dhillon, S., Lambert, P., Nachega, J. B., Wilmotte, L., et al. (2007). Effect of a collaborative approach on the quality of prescribing for geriatric inpatients: A randomized, controlled trial. *Journal of the American Geriatrics Society*, *55*(5), 658-665.

Stevenson, K. B., McMahon, J. W., Harris, J., Hillman, J. R., & Helgerson, S. D. (2000). Increasing pneumococcal vaccination rates among residents of long-term-care facilities: Provider-based improvement strategies implemented by peer-review organizations in four western states. *Infection Control and Hospital Epidemiology*, *21*(11), 705-710.

Stewart, S., Pearson, S., & Horowitz, J. D. (1998). Effects of a home-based intervention among patients with congestive heart failure discharged from acute hospital care. *Archives of Internal Medicine*, *158*(10), 1067-1072.

Stock, J. L., Waud, C. E., Coderre, J. A., Overdorf, J. H., Janikas, J. S., Heiniluoma, K. M., et al. (1998). Clinical reporting to primary care physicians leads to increased use and understanding of bone densitometry and affects the management of osteoporosis. *Annals of Internal Medicine*, *128*(12 Part 1), 996-999.

Szpunar, S. M., Williams, P. D., Dagroso, D., Enberg, R. N., & Chesney, J. D. (2006). Effects of the tobacco use cessation automated clinical practice guidelines. *American Journal of Managed Care*, *12*(11), 665-673.

Tay, S.-K., Tsakok, F. H., & Ng, C. S. (1992). The use of intradepartmental audit to contain cesarean section rate. *International Journal of Gynecology & Obstetrics*, *39*(2), 99-103.

Teich, J. M., Merchia, P. R., Schmiz, J. L., Kuperman, G. J., Spurr, C. D., & Bates, D. W. (2000). Effects of computerized physician order entry on prescribing practices. *Archives of Internal Medicine*, *160*(18), 2741-2747.

Thomas, M., Gillespie, W., Krauss, J., Harrison, S., Medeiros, R., Hawkins, M., et al. (2005). Focus group data as a tool in assessing effectiveness of a hand hygiene campaign. *American Journal of Infection Control*, *33*(6), 368-373.

Tran, T. P., Schutte, W. P., Muelleman, R. L., & Wadman, M. C. (2002). Provision of clinically based information improves patients' perceived length of stay and satisfaction with EP. *American Journal of Emergency Medicine*, *20*(6), 506-509.

van Kasteren, M. E., Mannien, J., Kullberg, B. J., de Boer, A. S., Nagelkerke, N. J., Ridderhof, M., et al. (2005). Quality improvement of surgical prophylaxis in Dutch hospitals: Evaluation of a multi-site intervention by time series analysis. *Journal of Antimicrobial Chemotherapy*, *56*(6), 1094-1102.

Verstappen, W. H., van der Weijden, T., Sijbrandij, J., Smeele, I., Hermsen, J., Grimshaw, J., et al. (2003). Effect of a practice-based strategy on test ordering performance of primary care physicians: A randomized trial. *Journal of the American Medical Association*, *289*(18), 2407-2412.

Wahlstrom, R., Kounnavong, S., Sisounthone, B., Phanyanouvong, A., Southammavong, T., Eriksson, B., et al. (1993). Effectiveness of feedback for improving case management of malaria, diarrhoea and pneumonia: A randomized controlled trial at provincial hospitals in Lao PDR. *Tropical Medicine and International Health*, *8*(10), 901-909.

Weller, T. M. (2002). The successful introduction of a programme to reduce the use of IV ciprofloxacin in hospital. *Journal of Antimicrobial Chemotherapy*, *49*(5), 827-830.

Winickoff, R. N., Coltin, K. L., Morgan, M. M., Buxbaum, R. C., & Barnett, G. O. (1984). Improving physician performance through peer comparison feedback. *Medical Care*, *22*(6), 527-534.

Wolverton, C. L., Hobbs, L. A., Beeson, T., Benjamin, M., Campbell, K., Forbes, C., et al. (2005). Nosocomial pressure ulcer rates in critical care: Performance improvement project. *Journal of Nursing Care Quality*, *20*(1), 56-62.

Won, S. P., Chou, H. C., Hsieh, W. S., Chen, C. Y., Huang, S. M., Tsou, K. I., et al. (2004). Handwashing program for the prevention of nosocomial infections in a neonatal intensive care unit. *Infection Control and Hospital Epidemiology*, *25*(9), 742-746.

(continued)

Appendix (continued)

Wyatt, J. C., Paterson-Brown, S., Johanson, R., Altman, D. G., Bradburn, M. J., & Fisk, N. M. (1998). Randomised trial of educational visits to enhance use of systematic reviews in 25 obstetric units. *British Medical Journal*, *317*(7165), 1041-1046.

Wyman, J. F., Croghan, C. F., Nachreiner, N. M., Gross, C. R., Stock, H. H., Talley, K., et al. (2007). Effectiveness of education and individualized counseling in reducing environmental hazards in the homes of community-dwelling older women. *Journal of the American Geriatrics Society*, *55*(10), 1548-1556.

Yoo, S., Ha, M., Choi, D., & Pai, H. (2001). Effectiveness of surveillance of central catheter-related bloodstream infection in an ICU in Korea. *Infection Control and Hospital Epidemiology*, *22*(7), 433-436.

Zack, J. E., Garrison, T., Trovillion, E., Clinkscale, D., Coopersmith, C. M., Fraser, V. J., et al. (2002). Effect of an education program aimed at reducing the occurrence of ventilator-associated pneumonia. *Critical Care Medicine*, *30*(11), 2407-2412.

Zanetti, G., Flanagan, H. L., Cohn, L. H., Giardina, R., & Platt, R. (2003). Improvement of intraoperative antibiotic prophylaxis in prolonged cardiac surgery by automated alerts in the operating room. *Infection Control and Hospital Epidemiology*, *24*(1), 13-16.

Zgibor, J., Orchard, T., Saul, M., Piatt, G., Ruppert, K., Stewart, A., et al. (2006). Developing and validating a diabetes database in a large health system**.** *Diabetes Research and Clinical Practice*, *75*(3), 313-319.

Zhu, S.-H., Tedeschi, G., Anderson, C. M., Rosbrook, B., Byrd, M., Johnson, C. E., et al. (2000). Telephone counseling as adjuvant treatment for nicotine replacement therapy in a "real-world" setting. *Preventive Medicine*, *31*(4), 357-363.

Zwar, N., Wolk, J., Gordon, J., Sanson-Fisher, R., & Kehoe, L. (1999). Influencing antibiotic prescribing in general practice: A trial of prescriber feedback and management guidelines. *Family Practice*, *16*(5), 495-500.

References

Barnes, R. V., Lawton, L., & Briggs, D. (1994). Clinical benchmarking improves clinical paths: Experience with coronary artery bypass grafting. *Joint Commission Journal on Quality Improvement*, *20*(5), 267-276.

Brook, R. H., Kamberg, C. J., Mayer-Oakes, A., Beers, M. H., Raube, K., & Steiner, A. (1990). Appropriateness of acute medicine care for the elderly: An analysis of the literature. *Health Policy*, *14*(3), 225-242.

Brothers, T. E., Robison, J. G., & Elliott, B. M. (1997). Relevance of quality improvement methods to surgical practice: Prospective assessment of carotid endarterectomy. *American Surgeon*, *63*(3), 213-219.

Carey, R. G., & Teeters, J. L. (1995). CQI case study: Reducing medication errors. *Joint Commission Journal on Quality Improvement*, *21*(5), 232-237.

Chu, L. A., Bratzler, D. W., Lewis, R. J., Murray, C., Moore, L., Shook, C., et al. (2003). Improving the quality of care for patients with pneumonia in very small hospitals. *Archives of Internal Medicine*, *163*(3), 326-332.

Civetta, A. C., Keller, R. B., & Wennberg, D. E. (1996). Decreasing catheter-related infection and hospital costs by continuous quality improvement. *Critical Care Medicine*, *24*(9), 1660-1665.

Daucourt, V., Saillour-Glenisson, F., Michel, P., Jutand, M. A., & Abouelfath, A. (2003). A multicenter cluster randomized controlled trial of strategies to improve thyroid function testing. *Medical Care*, *41*(3), 432-441.

Dellefield, M. E. (2006). Organizational correlates of the risk-adjusted pressure ulcer prevalence and subsequent survey deficiency citation in California nursing homes. *Research in Nursing and Health*, *29*, 345-358.

Deming, W. E. (1986). *Out of the crisis.* Cambridge: Massachusetts Institute of Technology Center for Advanced Engineering Study.

Dubois, R. W., & Brook, R. H. (1988). Preventable deaths: Who, how often, and why? *Annals of Internal Medicine*, *109*(7), 582-589.

Eccles, M., Grimshaw, J., Campbell, M., & Ramsay, C. (2003). Research designs for studies evaluating the effectiveness of change and improvement strategies. *Quality and Safety in Health Care*, *12*(1), 47-53.

Eccles, M., Steen, N., Grimshaw, J., Thomas, L., McNamee, P., Soutter, J., et al. (2001). Effect of audit and feedback, and reminder messages on primary-care radiology referrals: A randomised trial. *Lancet*, *357*(9266), 1406-1409.

Ellrodt, A. G., Conner, L., Rieidinger, M., & Weingarten, S. R. (1995). Measuring and improving physician compliance with clinical practice guidelines: A controlled intervention trial. *Annals of Internal Medicine*, *122*(4), 277-282.

Ferguson, T. B., Jr., Peterson, E. D., Coombs, L. P., Eiken, M. C., Carey, M. L., Grover, F. L., et al. (2003). Use of continuous quality improvement to increase use of process measures in patients undergoing coronary artery bypass graft surgery: A randomized controlled trial. *Journal of the American Medical Association*, *290*(1), 49-56.

Gregor, C., Pope, S., Werry, D., & Dodek, P. (1996). Reduced length of stay and improved appropriateness of care with a clinical path for total knee or hip arthroplasty. *Joint Commission Journal on Quality Improvement*, *22*(9), 617-627.

Grimshaw, J. M., Thomas, R. E., MacLennan, G., Fraser, C., Ramsay, C. R., Vale, L., et al. (2004). Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technology Assessment*, *8*(6), 1-94.

Grol, R., Baker, R., & Moss, F. (2002). Quality improvement research: Understanding the science of change in health care: Essential for all who want to improve health care. *Quality and Safety in Health Care*, *11*(2), 110-112.

Harris, R. H., MacKenzie, T. D., Leeman-Castillo, B., Corbett, K. K., Batal, H. A., Maselli, J. H., et al. (2003). Optimizing antibiotic prescribing for acute respiratory tract infections in an urban urgent care clinic. *Journal of General Internal Medicine*, *18*(5), 326-334.

Health Care Advisory Board. (1992). *TQM: Directory of hospital projects*. Washington, DC: Advisory Board Company.

Hoff, T., Jameson, L., Hannan, E., & Flink, E. (2004). A review of the literature examining linkages between organizational factors, medical errors, and patient safety. *Medical Care Research and Review*, *61*(1), 3-37.

Horbar, J. D., Carpenter, J. H., Buzas, J., Soll, R. F., Suresh, G., Bracken, M. B., et al. (2004). Collaborative quality improvement to promote evidence based surfactant for preterm infants: A cluster randomised trial. *British Medical Journal*, *329*(7473), 1004-1007.

Institute of Medicine. (2000). *To err is human: Building a safer health system*. Washington, DC: National Academy Press.

Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academy Press.

Ishikawa, K. (1985). *What is total quality control? The Japanese way.* (D. Lu, Trans.). Englewood Cliffs, NJ: Prentice Hall.

James, B. C. (1989). *Quality management for health care delivery.* Chicago: Hospital Research and Educational Trust.

Joshi, M. S., Anderson, J. F., & Marwaha, S. (2002). A systems approach to improving error reporting. *Journal of Healthcare Information Management*, *16*(1), 40-45.

Juran, J. M. (1989). *Leadership for quality.* New York: Free Press.

Leape, L. L. (1994). Error in medicine. *Journal of the American Medical Association, 272*(23), 35-43.

Marrie, T. J., Lau, C. Y., Wheeler, S. L., Wong, C. J., Vandervoort, M. K., Feagan, B. G., et al. (2000). A controlled trial of a critical pathway for treatment of community-acquired pneumonia. *Journal of the American Medical Association*, *283*(6), 749-755.

McLaughlin, C. P., & Kaluzny, A. D. (2006). *Continuous quality improvement in health care*. Sudbury, MA: Jones & Bartlett.

Meyer, G. S., & Massagli, M. P. (2001). The forgotten component of the quality triad: Can we still learn something from "structure"? *Joint Commission Journal on Quality and Patient Safety*, *27*(9), 484-493.

Moray, N. (1994). Error reduction as a systems problem. In M. S. Bogner (Ed.), *Human error in medicine* (pp. 67-91). Hillsdale, NJ: Lawrence Erlbaum.

Pandey, D., & Cursio, J. (2006). Data feedback for quality improvement of stroke care CAPTURE stroke experience. *American Journal of Preventive Medicine*, *31*(6), S224-S229.

Schectman, J. M., Schroth, W. S., Verme, D., & Voss, J. D. (2003). Randomized controlled trial of education and feedback for implementation of guidelines for acute low back pain. *Journal of General Internal Medicine*, *18*(10), 773-780.

Serisier, D. J., & Bowler, S. D. (2007). Effect of a simple educational intervention on the hospital management of community-acquired pneumonia. *Respirology*, *12*, 389-393.

Shojania, K. G., Yokoe, D., Platt, R., Fiskio, J., Ma'luf, N., & Bates, D. W. (1998). Reducing vancomycin use utilizing a computer guideline: Results of a randomized controlled trial. *Journal of the American Medical Informatics Association*, *5*(6), 554-562.

Shortell, S. M., Bennett, C. L., & Byck, G. R. (1998). Assessing the impact of continuous quality improvement on clinical practice: What it will take to accelerate progress. *Milbank Quarterly*, *76*(4), 593-624.

Szpunar, S. M., Williams, P. D., Dagroso, D., Enberg, R. N., & Chesney, J. D. (2006). Effects of the tobacco use cessation automated clinical practice guidelines. *American Journal of Managed Care*, *12*(11), 665-673.

Walton, M. (1990). *Deming management at work.* New York: G. P. Putnam.