Keys for successful implementation of total quality management in hospitals

James M. Carman
Stephen M. Shortell
Richard W. Foster
Edward F.X. Hughes
Heidi Boerstler
James L. O' Brien
Edward J. O'Connor

Editor's Note: This article reports the findings of an analysis of the implementation of continuous quality improvement (CQI) or total quality management (TQM) programs in 10 hospitals. This analysis is the result of a 2-year study designed to identify and assess the ingredients that lead to the successful implementation of CQI programs in acute care hospitals. This article first appeared in *Health Care Management Review 21*(1), 48–60.

In recent years, there has been significant interest in the application of continuous quality improvement (CQI) into health care organizations around the globe. In the United States, a Baldrige Quality Award specific to health care is planned for 1996. The hospital industry, in particular, has substantially embraced the concepts of CQI and total quality management (TQM) with the belief that these concepts and programs will lead to an improvement in both the quality and efficiency with which health services are delivered. Here we use these terms synonymously to mean an ongoing effort to provide services that meet or exceed customer expectations through a structured

Key words: Baldrige Award, CQI, health care management, TQM

James M. Carman, BS, MBA, PhD, *is Professor of Business Administration, Haas School of Business, University of California at Berkeley.*
Stephen M. Shortell, PhD, *is A.C. Burehler Distinguished Professor of Health Services Management, J. L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL.*
Richard W. Foster, PhD, *is Associate Professor, Graduate School of Business Administration, University of Colorado at Denver.*
Edward F.X. Hughes, MD, MPH, *is Professor, J. L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL.*
Heidi Boerstler, PhD, *is Associate Professor, Graduate School of Business Administration, University of Colorado at Denver, Denver, CO.*
James L. O'Brien, MA, *is Senior Research Associate, Center for Health Services and Policy Research, Northwestern University, Evanston, IL.*
Edward J. O'Connor, PhD*, is Professor, Graduate School of Business Administration, University of Colorado at Denver, Denver, CO.*

The research upon which this article is based was supported by the Center for Health Management Research of the Network of Healthcare Management and the National Science Foundation. Appreciation is expressed to the following participating institutions: Franciscan health System Group West (Tacoma), Good Samaritan Regional Medical Center (Phoenix), Intermountain Health Care, Inc. (Salt Lake City), Mercy Health Services (Farmington Hills, MI), Sisters of Charity Health Care Systems, Inc. (Cincinnati), Sisters of Providence (Seattle), Tucson Medical Center (Tucson), and Virginia Mason Medical Center (Seattle). Special acknowledgment is due the members of the research advisory group of the participating institutions for their advice and support throughout the study. The research assistance of Judith Strebel of the Haas School of Business, University of California at Berkeley is acknowledged and greatly appreciated.

DOI: 10.1097/HMR.0b013e3181f5fc4a

Health Care Manage Rev, 2010, 35(4), 283-293

systematic process for creating organizationwide participation in planning and implementing quality improvements. This process includes: (a) an organizational structure for identifying and improving processes, (b) use of a set of data-based statistical and analytical tools to study processes, (c) empowerment of teams of employees to take charge of the operations of their own work tasks in a manner that encourages continuous learning as well as empowerment and personal responsibility. It requires that management leadership create an organizational culture committed to continuous improvement and learning as opposed to merely correcting deficiencies or meeting current standards.

Building on our prior work, which examined the relationships among QI implementation approach, culture, and clinical outcomes, (Shortell et al., 1995) the present article focuses on the findings derived from extensive study and site visits to 10 hospitals.

Study Framework and Hypotheses

The role of culture

The overall framework for the analysis reported here is shown in Figure 1. This framework places more emphasis than was true in Shortell et al. (1995) on the key role played by organizational culture and on outputs and performance measured in global terms rather than specific clinical conditions.

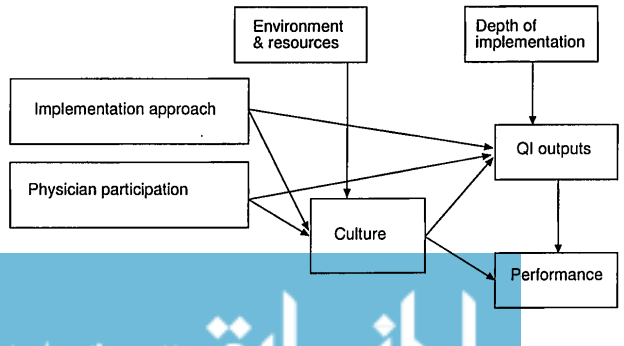
Culture is hypothesized to be an intervening variable between the implementation approach and the other causal factors. In other words, these other factors are hypothesized to work on outputs in two ways, that is to say, directly on outputs and also indirectly on outputs and performance through culture. Placing culture in this pivotal role stems from the overriding importance of culture for successful implementation of CQI observed in the site visits and demonstrated by the empirical evidence reported in the earlier article. Two key hypotheses are:

H1: An empowered and continuous learning culture among members of the work force will create greater quality outputs and performance that exceed the effects of team process improvement alone.

H2: The performance improvement results from a CQI program can be seen in customer satisfaction and market share as well as economic efficiency as measured by length of stay, costs per service, and labor productivity.

Figure 1

Conceptual model for assessing the impact of quality management



Environment and resources

In the earlier analysis we found that hospital size, as measured by bed size, had an influence on culture, implementation success, and clinical efficiency. In the present analysis the measurement of hospital size and complexity is enriched and other environmental factors concerning competitive pressures are added. Specifically,

H3: Smaller hospitals with fewer complex services will more easily implement a CQI program than will large hospitals with a multiplicity of inpatient and outpatient services generating greater revenue.

H4: Hospitals whose staff associates perceive an intense competitive environment will experience more implementation success than will hospitals whose staffs do not feel economically threatened.

H5: Hospitals that have taken on more economic risk in terms of capitated contracts will experience more implementation success than will hospitals that have not taken on capitated business.

Implementation approach

Two constructs for describing implementation approach were employed in the study on which this article is based. One was a variant on the Miles and Snow (1978) strategic style construct. Meyer, Brooks, and Goes (1990) and Shortell and Zajac (1990) have applied this typology to competitive strategies among hospitals, but this was its first application to implementing CQI. The Miles and Snow typology involves classifying organizations as analyzers, prospectors, defenders, and reactors. Analyzers develop a new program by following an established sequence of steps. Prospectors consciously attempt to be more innovative in new program development. Defenders fine tune existing programs. Reactors change only as forced by external pressures. These four types are defined in more detail in Shortell et al. (1995). CQI is such a major departure from traditional management

philosophies that we believe marginal alterations to traditional quality assurance will not be successful. Thus the hypothesis that:

> *H6:* Hospitals using analyzer and prospector implementation approaches will experience a greater degree of CQI implementation than those using defender or reactor approaches.

The second implementation approach construct was suggested by Berwick, Godfrey, and Roessner (1990). This construct classifies programs by the emphasis or motivation for the program in the first years. There are three classes: a project dominant emphasis that identifies key processes that need to be addressed (similar to a reengineering approach); a strategy dominant emphasis that identifies quality as the competitive strategy of the hospital required to ensure survival, market share, and please third party payers; a culture dominant emphasis that identifies vision, values, and an empowered work force as the key objectives of the CQI program. The hypothesis is that:

> *H7:* Hospitals initially emphasizing a project dominant or strategy dominant approach will experience a greater degree of CQI implementation success than those emphasizing culture dominance.

The reasoning that lies behind this hypothesis is that changing the culture of the individuals in a work force is not something that can be done directly. Individual value systems and visions about self-worth and self-development change slowly if at all. Thus, the CQI approach is not to make a frontal attack on culture but rather to make a frontal attack on process improvement by associates in the work force. It is the experience of participation in such processes and training for process improvements that leads to individual improvement.

Physician participation

The traditional organizational structure of acute medical care in this country with physicians as independent providers and not hospital employees has meant that hospitals have found it difficult to directly involve physicians in their CQI programs. Hospitals have taken different approaches to dealing with this structural difficulty. For example, processes identified for study can be logically divided into two types, clinical and administrative. Some hospitals have introduced their CQI programs with administrative projects and teams only; others have begun their programs through physician "champions" who viewed CQI entirely in terms of clinical processes; still others have tried to balance these two approaches by encouraging a relatively small group of physician "champions" to join with administrators in agreeing to focus on an initial set of projects that involved both clinical and administrative processes. It was believed that the extent and methods for engaging physicians in the CQI program were so central to success that multiple measures of these characteristics were collected. Thus in Figure 1, physician participation is emphasized by identifying it as a separate block from implementation approach. Our hypo thesis was that:

> *H8:* Hospitals that successfully involve physicians in the CQI program early in the program will experience more implementation success than will hospitals that move ahead without significant physician participation.

Depth of implementation

The sample of hospitals for site visits was selected so as to provide variance in the length of time the hospitals had been involved in their CQI program. Roughly, two hospitals were just beginning; two had completed 1 year; three had completed 2 years; three had been involved in a formal CQI program for more than 3 years. Clearly, results would be expected to be related to the relative duration of the program. This variable is something more than self-evident because of the frequent criticism in the literature that CQI takes more than 4 years to produce results and therefore is too great an investment for many firms (Boerstler et al., 1996; Cole, 1995; Joint Commission on Accreditation of Healthcare Organizations, 1992). However, we found some organizations consciously chose to implement at a slower pace than others. Chronological time does not entirely measure depth. This phenomenon is reflected in the measures of depth used for this construct: One was based on the total number of CQI activities, and the other was related to the length of time the hospital had been involved in outcomes studies. Thus the hypothesis that:

> *H9:* QI output success will be positively correlated with the depth of involvement in the CQI program.

Data and Method

Sample

Site visit hospitals were selected through discussion with the eight systems listed in the acknowledgments section of this article. Two hospitals in the same system were selected in two cases, thus a sample of 10 hospitals. As indicated just above, one criterion used in selection was the length of time the hospital had been involved in its CQI program. Geographically, the hospitals are located in Michigan and westward with 6 of the 10 being in the Pacific time zone. The hospitals ranged in number of acute beds from 60 to 571 with mean of 360 and median

of 400. Sample hospitals are above national norms in terms of residency and research activities.

Between two and five senior investigators, mainly the authors, spent approximately 2% days at each site. The first half day was devoted to collection of additional secondary data. A minimum of 1% days was devoted to one-on-one interviews with hospital staff. The research team identified in advance the people to be interviewed in terms of their function, for example, chief executive, physicians involved in QI, team members, nursing director, quality manager. Each interviewer followed an interview guide designed for the respondent in each particular role. All interviews were recorded. The investigators developed an outline for summarizing impressions of the interviews. Very shortly after the site visits, all interviewers summarized their findings in writing. After a conference telephone call, these written reports were revised and sent to one member of the team who wrote a final report. Finally, a structured instrument with numerical evaluations of key dimensions was completed by each site visitor. Thus, site visitors before completing the numerical ratings of key variables had as background: their own impressions; preliminary discussions with other team members; their own written reports; additional discussions with other team members; verbal reaction from the hospital's management; a team written report; and considerable statistical data. Table 1 summarizes the measures tested in the analysis.

Measures

Environment and resources. Eight measures of the hospitals' environment and resources were tested. All of

Table 1

Description of variables

Variable description	Mean	Range
Environment and resources		
Annual gross revenue, millions of $	174.	21–410
Percent operating margin, % of gross	2.7	−1.0–12.6
Market share, %	24.7	7–57
Length of stay, days (LOS)	5.1	3.6–7.7
Cost/adj. admission, $	5,288.	3,600–6,909
Productivity, FTE/1000 adj. adm.	10.3	4.5–16.0
SV (site visitors) perceived market competitiveness, 5-point scale	3.6	1–5
3-year growth in percentage of capitated contracts	1.0	0–2.6
Implementation approach		
Analyzer FS (factor score)	0	−1.75–1.07
Prospector FS	0	−0.88–1.84
SV project dominant, dummy	.50	
SV strategy dominant, dummy	.20	
Physician participation		
FS of MDs in mgt. & SV MD participation	0	−1.28–1.63
FS of MDs reported participation in QI	0	−1.42–1.57
SV clinical emphasis, dummy	.40	
Culture		
Group culture score	27.8	20.9–38.4
Hierarchical culture score	29.4	22.3–36.4
Developmental culture score	17.9	15.3–21.0
SV strength of culture, 5-point scale	3.8	3–5
Role conflict FS: role conflict, supervisor support, distressing environment	0	−1.27–1.53
Depth of implementation		
Total volume of QI activities FS	0	−1.23–1.53
Use of outcomes & pat. sat. studies FS	0	1.50–1.01
QI outputs		
Bald. Quality results score, 5-point scale	3.5	3.2–3.7
Bald. customer focus score, 5-point scale	3.5	3.2–3.8
Sum of 6 Bald. input scores	19.8	18.0–20.9
SV total Bald. score, 0–100	38.5	10.9–62.9
Performance		
1-year change in pat. sat. scores, standardized	3.9	−3.82–17.13
3-year change in efficiency FS: LOS, cost/adm., productivity	0	−2.20–1.37
Change in percentage market share	9.2	−9.0–30.0

these variables were collected or verified for accuracy during or after the site visit. Three of the variables were market share, gross revenue, and percent operating margin. Three were efficiency measures: length of stay, cost per admission, and labor productivity. Two were concerned with competitiveness: site visitors' perceived rating of market competitiveness, and the growth rate of capitated, at-risk reimbursement contracts.

Implementation approach. Two conceptual typologies were tested in this research. One was based on the work in strategic management of Miles and Snow (1978). This typology involves classifying organizations as prospectors, analyzers, defenders, and reactors. The measurement scheme used here is a modification of the original methodology that had been employed by Shortell and Zajac (1990). A category of "beginners" was added to accommodate hospitals that had not yet developed an implementation strategy. Self-administered questionnaires concerning the hospital's approach to implementation were completed by senior executives, quality improvement council members, and quality managers.

Responses were then used to score hospitals on implementation approach. For this analysis, the scores were factor analyzed and hospitals were assigned continuous factor scores for three factors: analyzers, prospectors, and beginners. (Defender and reactor scores were not robust in the factor analysis.) An analyzer would attempt to maintain a relatively stable set of quality improvement activities for selected departments and conditions and would usually not be first to implement new activities. A prospector would emphasize frequent changes in the mix of quality improvement activities undertaken and would attempt to be first in implementing new activities.

The second typology of implementation approach was Berwick's project dominant, strategy dominant, or culture dominant approach (Berwick et al., 1990). This classification was made by site visitors after the site visit reports had been completed. Site visitors assigned each hospital to one of these three classifications both in terms of initial approach and, for those hospitals further along, approach after a trial period of a year or more. Again, hospitals just beginning their CQI program were simply classified as beginners. These assignments did not prove easy for site visitors even when they were permitted to classify a hospital as a combination type, for example, three hospitals were classified as having evolved into being "project/strategy dominant." The "farther along" classifications proved more useful than the initial classifications. For the purposes of the present analysis, these classifications were assigned to three dummy variables: project dominant, strategy dominant, or beginners.

Physician participation. Five measures of physician involvement in the CQI program were tested in some way in this analysis. They came from a variety of sources: the self-administered questionnaires completed by senior executives; quality improvement council members and quality managers; the employee perceptions questionnaire; and site visitor ratings. These measures concerned the extent of physician involvement in hospital management, site visitor perceptions of physician involvement, employee perceptions of physician leadership in the CQI program, and the number of teams with physician participation. These four were simplified into two factor scores. The fifth measure, clinical emphasis, was a dummy variable based on classification by site visitors of whether the hospital had clinical emphasis, administrative emphasis, or balanced emphasis in its projects and approach to CQI implementation.

Culture. Organizational culture has emerged as a widely studied phenomenon in management research. It was not obvious in the planning of this study just which measure of workplace culture would be most appropriate. A number were tried.

One of these, based on the work of Kimberly and Quin (1984) had been tested in a pilot study in three hospitals and appeared to be a powerful measure. It proved to be the most important measure. Kimberly and Quinn defined four cultural types: a *group culture* based on norms and values associated with affiliation and teamwork; *a developmental culture* based on assumptions of change and risk-taking; a *hierarchical culture* reflecting the values and norms associated with bureaucracy such as control, stability, and security; and a *rational culture* emphasizing productivity and efficiency. An organization is likely to exhibit some characteristics of all four types. The question is which dominates the organization's value system. Hypothesis 1 speaks of an empowered and continuous learning culture. It is believed that these values would be positively associated with a *group* or *developmental* culture and negatively associated with a hierarchical culture. The emphasis placed on each of these four types as perceived by employees were measured with a 20-item scale developed by Zammuto and Krakower (1991) and pretested in nonstudy hospitals. The reliability measures for three of these four scales had alphas between .70 and .79, while the alpha for the rational scale was .47 (Shortell et al., 1995). Senior executives and quality council members completed this instrument.

Another culture measure came from site visitors who rated on a five-point scale the strength or intensity of the hospital's culture, whatever type it may have been.

A final measure of relating to workplace culture was a role conflict score composed of three elements. One was the well-known employee perceptions of role conflict scale of Rizzo, House, and Lirtzman (1970) and Schneider, Parkington, and Buxton (1980) but modified to fit the hospital setting. This scale had been tested previously in a hospital setting and had a reliability

alpha of .77. The second element was a 21-item scale measuring employee perceptions of supervisor support. Again, this scale had been tested previously in a hospital setting and had a reliability alpha of .97. The third element was a measure of employee feeling of distress in the workplace. Some of these had been significant in predicting employee perceptions of CQI success when the individual employee, rather than the hospital, was used as the unit of analysis.

Depth of implementation. Two scales were used to measure the depth of implementation of the hospital in the CQI program. Note that Hypothesis 9 is not expressed in terms of length of time involved in a CQI program. While this is one measure, and one used here, it may not be the best measure. We found in the site visits that hospitals differed in how aggressively they introduced CQI. Some hospitals spent a year or more in behind-the-scenes planning. Thus, the measures used here were designed to measure depth of implementation in ways other than just chronological time.

One measure was a combination of scales that counted the number of CQI elements and the volume of activity in each element. For example, having an inhouse quality training program was one CQI element. Counting the number of teams actively working was a volume of activity scale.

The second measure related more to chronology in that it was based on the length of time that clinical performance studies had been underway, and the extent of the link between the CQI program and patient satisfaction monitoring. All of the data for these measures came from baseline data collection done by each hospital prior to the site visits.

QI outputs. There are two sources of measures of quality outputs. One was from employee perceptions; the other was from site visitor perceptions. In both cases, outputs were defined in terms of the dimensions of the Malcolm Baldrige National Quality Award Criteria (U.S. Chamber of Commerce, 1993). In the case of site visitors, the mean of the Total Baldrige Score was used. For employees, three components of the Baldrige score were used. These components will be defined shortly.

The employee Baldrige questionnaire was specially designed for this study so as to be applicable to hospital settings. Items measuring the extent of progress in a TQM program were tested by administration to employees of three nonstudy hospitals and revised after that test. The instrument contained 76 items that factored into 8 dimensions. The reliability tests of these scales produced alphas from .79 to .93 (Shortell et al., 1995).

Our conceptualization of the Baldrige schema involves thinking of the eight dimensions as six input measures [(1) leadership, (2) strategic quality planning, (3) education, (4) empowerment, (5) information and analysis employed, and (6) management of the quality improvement process] and two output measures [(7) quality results and (8) customer satisfaction]. In the empirical study discussed here, one QI output measure was an overall measure of CQI inputs constructed by summing scales 1 through 6 (the simple correlations of these six with the summed measure ranged from .83 to .97); another was (7) the quality results score; a third was (8) the customer satisfaction score. In all cases, the mean of the employee sample was used as the hospital's score. It should be emphasized that these are considered output measures and not performance measures because they represent only employee perceptions, not customer perceptions or observable facts.

Performance. The CQI program may be expected to improve performance in terms of increased economic efficiency (i.e., length of stay, costs, and labor productivity), improved clinical outcomes, improved customer satisfaction, and increased market acceptance.

> The change in customer satisfaction required dealing with patient satisfaction surveys that were not comparable.

The change in overall hospital length of stay was calculated over a 3-year period ending with the site visit year. (Note that length of stay was used as an efficiency measure: *change in length of stay* as a performance measure.) The change in labor productivity was measured as the change in full-time equivalency (FTE) personnel per adjusted admission also over a 3-year period. The change in adjusted cost per admission and the change in market share were calculated over a 2-year period. All of these measures were readily available in relatively comparable form from the hospitals.

The change in customer satisfaction required dealing with past patient satisfaction surveys that were not comparable. Questions dealing with particular services were eliminated, so only overall evaluations were considered. Then the scores were standardized so as to have comparable mean and variance across hospitals. Finally, the most recent 1-year change in past patient satisfaction score was calculated. It was not possible to take a longer time interval because hospitals had changed questionnaires or had not conducted past patient surveys before the present decade.

Method

The process of testing proceeded as follows. First, multiple regressions of each path in Figure 1 were run and clearly unimportant variables were eliminated. Then ordinary least squares or, where appropriate, two-stage least squares regressions were run on sections of the model that are largely independent of one another. To be precise, a two-stage least squares model with hierarchical culture was run

as one regression, while a two-stage least squares regression model with strength of culture were run independently. While this methodology would not be appropriate if the sample of over 100 hospitals were available, the predictors left in the model were quite robust and provide a useful basis for testing the hypotheses.

Findings

The final version of the conceptual model along with the estimated standardized regression coefficients for each path are shown in Figure 2. The findings for each section are now described.

Environment and resources-the effect of size of facility on culture

Despite the fact that measures of profitability and the pressure to shift risk to providers were tested as environmental drivers to successful CQI implementation, the only two of the eight variables relating to environment and resources tested that were significant were gross revenue and length of stay (LOS). These variables are really proxies for size and complexity of the hospital. Both had very strong relationships with a hierarchical culture. In other words, by virtue of their size, large hospitals with subspecialty services tend to be more hierarchical than smaller, community hospitals. It is interesting to note that in the modern environment, bed size does not necessarily have a perfect correlation with volume or sophistication. Thus, in this analysis, gross revenue and LOS were better predictors than the number of beds.

Among the other environmental predictors tested were a measure of site visitors rating of the competitiveness of the market and the increase over the past 3 years in the number of patients covered by capitated contracts. Neither of these variables was significant in the final model.

Implementation approach

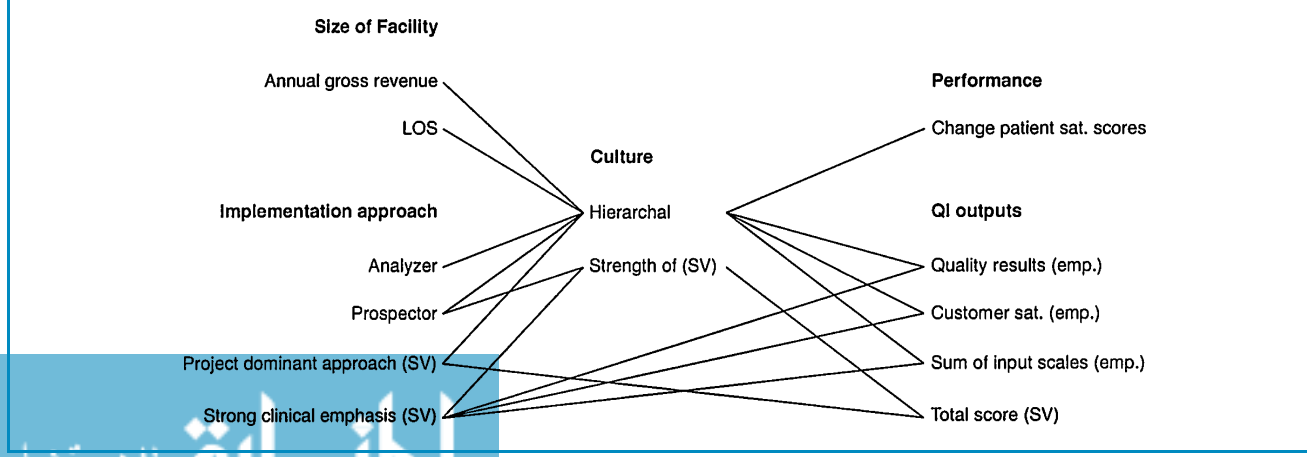
As described earlier, a major descriptor of implementation approach used in this study was a variant of the Miles and Snow (1978) typology. Indeed, in this analysis, a variety of forms of our measures were tested. Shortell et al. (1995) found that hospitals employing a prospector approach had enjoyed greater success in QI implementation (the sum of input scales employee perception measure in this article). The present analysis does not support this finding. The prospector and analyzer implementation approaches had a weak positive relationship with hierarchical cultures, and hierarchical culture had a negative impact on sum of input scales. Furthermore, prospectors were associated with weak cultural strength, while strong cultures had a positive impact on QI outputs.

The other descriptor of implementation approach classified hospitals into using a project, strategy, or culture dominant approach. Site visitors did not have an easy time agreeing on a single dominant approach. Project dominance often was mixed with, or transitioned into, a culture or strategy approach. In any case, hospitals that started with a project dominant approach had a negative relationship with hierarchical culture, and project dominance had a positive effect on QI outputs in addition to its effect on QI outputs through culture.

Physician participation. Three measures were used to test the hypothesis that QI success was improved by involving physicians in the CQI program. Two of these were

Figure 2

Path coefficients for final estimated model



the factor scores described above. The third measure was a site visitor judged dummy variable of the extent to which the QI program had a clinical emphasis. Of these three, only the third was significant in the final model. Hospitals judged by site visitors as having a strong clinical emphasis were perceived as having a strong culture and were perceived by employees as having achieved more success with their quality program.

Culture

All four of the Kimberly and Quinn (1984) culture types were tested in the path model. All had similar patterns of correlations with the predictors. However, only group culture and hierarchical culture had significant paths to QI outputs or performance. Of these two, the negative effects of a hierarchical culture on QI outputs and performance were somewhat stronger. It is the culture type that is shown in Figure 2.

The other culture variable left in the final model was the site visitor rating of strength of culture. This variable had a positive relationship with site visitors' total score on the Baldrige output dimensions. The role conflict variable was not significant when the unit of analysis was the hospital as it had been when the individual respondent was the unit of analysis.

Correlates with group culture

In the final estimated model, only the hierarchical culture score was left in the model. This should not be interpreted as suggesting that the other dimensions of the Kimberly and Quinn (1984) culture typology were not important. In this section, the findings in this area will be described in a bit more detail. As described in Shortell et al. (1995), this is a scale in which respondents must have their total score sum to 100. Thus, when one scale received a very large number of points, others must necessarily get a small number of points. In this study, developmental cultures and rational cultures received a relatively small number of points while most employees described their workplace culture as either group or hierarchical. Probably because of their low scores and some reliability concerns with the developmental and rational scales, these two scores did not have significant effects on QI output scores.

Employees viewed group and hierarchical cultures as opposites. If group culture received a large number of points, hierarchical received a small number. In Shortell et al. (1995), the group and developmental scores were combined and that sum score had *the* most important, and *positive*, effect upon implementation success. In the present analysis, the hierarchical score was found to have *the* most important, and *negative*, effect upon implementation success. When group culture is substituted for hierarchical culture in Figure 2, the relative importance of the paths are similar with opposite signs. In other words, these two are mirror image measures of one another. This comparison is shown in Table 2.

Depth of implementation

As shown in the conceptual model, it was hypothesized that depth of implementation, for example, how deeply the hospital was involved in CQI program, would influence the degree of success. While not shown in the final model, this hypothesis could not be rejected as a result of the empirical analysis. The second measure, based on length of involvement in clinical process studies and CQI influences on customer satisfaction, did have a significant influence on all three of the employee perception Baldrige measures shown in Figure 2. However, these measures were not included in the final estimated model because they swamped the effects of other variables on outputs and performance. It seemed more insightful to demonstrate the effects of these other factors thought to cause positive results from a CQI program.

To expand on this point, the two measures explained 61 percent of the variance in sum of input scales, and in that model, no other predictors were significant. The two paths shown in Figure 2 explained 81 percent of the variance in sum of input scales. With respect to the customer satisfaction dependent variable, the second measure of depth of implementation explained 37 percent of the residual variance after the effects of the two paths shown in Figure 2 were accounted for. However, these variables were not significant in explaining the residual variance in the other measures of QI output.

Table 2

Comparison of path coefficients for hierarchical and group culture scores

	Path coefficients from/to	
Variables	Hierarchical culture	Group culture
Predecessors		
Annual gross revenue	.83	−.60
Length of stay	.76	−.63
Analyzer	ns	ns
Prospector	.43	−.63
Project dominant	−.47	.33
Successors		
Change in pat. sat.	−.93	ns
Bald. quality results	−.44	.55
Bald. customer sat.	−.71	.70
Sum of input scales	−.51	.49

Effects on performance

As shown in Figure 1, the conceptual model hypothesized that successful QI outputs would have a positive influence on market performance of the hospital. In addition, it was hypothesized that culture would have an additional, direct effect on performance. Three measures of performance were tested: the 1-year change in overall past patient satisfaction scores; the 2-year change in market share; and a factor score measuring the change in economic efficiency as measured by overall LOS, adjusted cost per admission, and labor productivity.

The results were somewhat disappointing. None of the QI output measures influenced these performance variables. Hierarchical culture had the expected negative, and very strong, influence on change in patient satisfaction scores. However, it did not have a significant influence on the other two measures of performance. More will be said regarding this finding in the next section.

Discussion

The summary discussion of these findings will be organized principally around the nine hypo theses along with some discussion of measurement and validity matters.

H1: The first hypothesis stated that an empowered work force will create greater outcome gains than the effects of specific team process improvement alone. The hypothesis is supported in that, in terms of both employee and site visitor evaluations, a group, nonhierarchical culture not only served as an intervening variable between the elements of implementation but, in the case of clinical emphasis, had an impact on outcomes in addition to the direct impact of a clinical emphasis.

H2: At least in the time span of this study, the influence of culture and QI outputs on overall performance was not as great as that stated in the hypothesis. The Baldrige scores had no impact on performance, and culture influenced patient satisfaction scores but not market share or economic efficiency.

There may be a number of reasons why these effects were not stronger. One is that the time periods of change in performance did not correspond well enough to the time period of CQI activity. A second is that CQI had not had sufficient time to influence economic performance. With the environmental influences on economic performance so strong in health care today, a successful CQI program should permit survival and some improvement in efficiency and market share. However, such effects are probably not going to be evident in the first year or two of the program. Indeed a third reason is that in the short time period of this study, the effects of CQI were probably swamped by other environmental impacts, such as competition, that were influencing the hospitals.

A fourth reason, that can be seen in Shortell et al. (Shortell et al., 1995), is that the measures of performance used here were too aggregated. In other words, to see CQI impact after the first year or two requires measuring specific areas where CQI activities have been completed rather than looking at overall performance.

H3: The third hypothesis, which said that smaller, general hospitals with fewer tertiary services and teaching programs would have an easier time in implementing CQI than would more complex hospitals, was supported. Our findings reinforce a corollary that the number of acute beds operated is not necessarily the richest available measure of complexity.

H4 and H5: These two hypotheses posited that hospital managements and staffs that felt threatened by the environment or had taken on a substantial number of at-risk contracts would more easily implement their CQI program than would staffs that did not feel such threats. At least with the measures employed here, these hypo theses were not supported.

H6: This hypothesis posited that hospitals using an analyzer or prospector implementation approaches would experience a greater degree of CQI implementation than those using other approaches. This hypothesis was supported by our analysis of the 61 hospitals reported in our earlier article. Here the results were reversed. Prospectors were significant and analyzers mildly significant but with the wrong signs.

It is not clear to us whether the difference in findings was due to weak measures of the Miles and Snow (1978) typology, which is our present feeling, or whether this typology is not usefully applied to descriptions of CQI implementation approach. It is true that site visitors had trouble in classifying hospitals within this typology. The matter deserves more attention in future research.

H7: This hypothesis also concerned implementation approach and posited that emphasizing a project dominant or strategy dominant approach would achieve greatest CQI implementation success. Actually, it was found that starting with a project dominant approach and transitioning to a strategy or cultural approach produced the greatest QI outputs. Further study of the hospitals in this sample would help to develop more insights into the precise methods of making this transition.

H8: This hypo thesis posited that early involvement of physicians in the CQI program would have positive effects on implementation success. While only one of the multiple measures of physician participation was

significant, we believe there was support for this hypothesis. The nonsignificant measures were largely concerned with organization structure characteristics. The significant factor was based on the site visit interviews in which investigators could really observe the extent of physician involvement in the program. While the message here is that more attention needs to be paid to measurement issues, there is no doubt in the minds of the researchers that physician involvement on clinical process improvement teams early in the program is essential for CQI success.

H9: The last hypothesis was aimed at getting at the issue of the time required to achieve CQI results. It used the term "depth of involvement" to suggest that it was not only chronological time but extent of organizational immersion in the program that is important. As explained in the findings section, the measure of depth of involvement had very strong significant influence on both the measures of QI outputs and on customer satisfaction that are not shown in Figure 2.

In sum, support was found for five of the nine hypotheses; one other found weak support in the data. One should not conclude from these findings that any of the nine hypotheses should be rejected. Rather it is that some measures need to be strengthened and that longer longitudinal study of these same hospitals is required to observe the full power of a quality improvement program.

Some Final Thoughts

CQI versus reengineering

The strong relationship between culture and the change in customer satisfaction scores suggests some lessons for successful management. However, the evidence in this study alone is certainly not sufficient to constitute proof that a work force that perceives itself as empowered and customer service oriented is *the* key to customer satisfaction. The findings here also suggest that an enthusiastic work force does not necessarily imply improvement in economic efficiency. However, combining the evidence from our two articles suggests that process improvements lead to efficiency improvements and that process improvements and an empowered work force produce improved customer satisfaction *and* improved efficiency.

These findings suggest a distinction between the concepts of reengineering and continuous quality improvement that is important and merits further study. This distinction may be hypothesized as follows.

Reengineering may produce process improvements without producing a change in workplace culture. Such gains may or may not be held and may or may not have a positive effect on customer satisfaction. All of the elements of continuous quality improvement taken together are designed to achieve a change in work force culture that will produce process improvements, facilitate holding the gains from process improvements, and impact favorably on customer satisfaction as well as increase economic efficiency.

Comparison of site visitor and employee perceptions of quality results

One aspect of this article, as contrasted with our earlier article, was the use of site visitor evaluations in addition to employee evaluations of the CQI program. The measures used by both site visitors and employees are based on the Baldrige Award Criteria. Probably the most important validity finding is the general agreement between the perceptions of site visitors and employees. The total scores have a correlation of .64 that is just significant with alpha risk of .05. There are three individual scales with low correlations: customer satisfaction, employee training, and employee empowerment.

These differences in perceptions are insightful. It appears both groups were concerned about elements of the program they judged to be most important or closest to their own interests. Site visitors were concerned about performance results and information systems support; employees were focused on training and empowerment. Both employees and external experts can fruitfully contribute to the evaluation of the success of a CQI program, but in addition, future research needs to focus more on objective measures of clinical outcomes, customer satisfaction, and economic performance.

Implications for research design

The findings reported here need to be considered within the context of the research design. We have tested 9 hypotheses employing over 30 variables with just 10 degrees of freedom. Thus, there must be reservation about power and representativeness. However, given the exploratory nature of a study whose purpose was to investigate a new practice in health care delivery management, the design was appropriate. A very large, representative sample would have created such problems of overaggregation that it is unlikely the implications for management, theory, or research would have been as great as they are here. Likewise, a few case studies that went into more detail than reported here would have generated even greater concerns about projectability to other institutions. In short, the present research design involving statistical analysis of a small sample is probably the ideal research design for the purposes of understanding and helping with new practices. The implications for success uncovered here are projectable to new health care settings.

The time span of the study also merits comment. The investigators and some of the hospitals involved feel a longitudinal study of greater than 2 years would be even more insightful. More needs to be understood about the role of culture change as an intervening variable that leads to improved performance.

In addition to the contribution to practice, the study has provided contributions for further research. Many sections of the article have cautioned about measurement concerns. Researchers in this area should study these carefully in designing studies of CQI programs. For example, at least in the early years, the unit of analysis might be a unit smaller than a hospital, for example, a service line. At the same time as the industry moves to integrated delivery systems, CQI programs need to encompass pre- and postacute care and CQI programs need to be integrated across these internal customer and supplier provider groups. We believe this study has laid a solid foundation for improving management practice and research in health care quality improvement around the globe.

References

Berwick, D. M., Godfrey, A. B., & Roessner, J. (1990). *Curing Health Care*. San Francisco: Jossey-Bass.

Boerstler, J., Foster, W., O'Connor, E. J., O'Brien, J. L., Shortell, S., Carman, J. M., et al. (1996). "Implementation of Total Quality Management: Conventional Wisdom Versus Reality." *Hospital and Health Services Administration*, *41*, forthcoming.

Cole, R. (1995). *The Death and Life of the American Quality Movement*. New York: Oxford University Press.

Joint Commission on Accreditation of Healthcare Organizations. (1992). *Striving Toward Improvement: Six Hospitals in Search of Quality*. Chicago: American Hospital Publishing.

Kimberly, J. R., & Quinn, R. E. (1984). *Managing Organizational Transitions*. Homewood, Ill.: Dow Jones-Irwin.

Meyer, A. D., Brooks, G., & Goes, J. B. (Summer 1990). "Environmental Jolts and Industry Revolution: Organizational Responses to Discontinuous Change." *Strategic Management Journal*, *11*, 93–110.

Miles, R. E., and Snow, C. C. (1978). *Organizational Strategy, Structure, and Process*. New York: McGraw-Hill.

Rizzo, J. R., House, R. J., & Lirtzman, S. I. (1970). "Role Conflict and Ambiguity in Complex Organizations." *Administrative Science Quarterly*, *15*, 150–63.

Schneider, B., Parkington, J. J., & Buxton, V. M. (June 1980). "Employee and Customer Perceptions of Service in Banks." *Administrative Science Quarterly*, *25*, 252–67.

Shortell, S. M., O'Brien, J. L., Carman, J. M., Foster, R. W., Hughes, E. F. X., Boerstler, H., et al. (1995). "Assessing the Impact of Continuous Quality Improvement/Total Quality Management: Concepts Versus Implementation." *Health Services Research*, forthcoming.

Shortell, S. M., & Zajac, E. J. (1990). "Perceptual and Archival Measures of Miles and Snow's Strategic Types: A Comprehensive Assessment of Reliability and Validity." *Academy of Management Journal*, *33*, 817–832.

US. Chamber of Commerce. (1993). *The Malcolm Baldrige Award*. 1992.

Zammuto, R. F., & Krakower, J. Y. (1991). "Quantitative and Qualitative Studies of Organizational Culture." *Research in Organizational Change and Development*. Greenwich, Conn.: JAI Press.