

Organizational Factors Associated with High Performance in Quality and Safety in Academic Medical Centers

Mark A. Keroack, MD, MPH, Barbara J. Youngberg, JD, MSW, Julie L. Cerese, MSN, Cathleen Krsek, MSN, MBA, Leslie W. Prellwitz, MBA, and Eoin W. Trevelyan, DBA

Abstract

Purpose

Leaders of academic medical centers (AMCs) are challenged to ensure consistent high performance in quality and safety across all clinical services. The authors sought to identify organizational factors associated with AMCs that stood out from their peers in a composite scoring system for quality and safety derived from patient-level data.

Method

A scoring method using measures of safety, mortality, clinical effectiveness, and equity of care was applied to discharge abstract data from 79 AMCs for 2003–2004. Six institutions (three top

and three average performers) were selected for site visits; the performance status of the six institutions was withheld from the site visit team. Through interviews and document review, the team sought to identify factors that were associated with the performance status of the institution.

Results

The scoring system discriminated performance among the 79 AMCs in a clinically meaningful way. For example, the transition of a typical 500-bed hospital from average to top levels of performance could result in 150 fewer deaths per year. Abstraction of key

findings from the interview notes revealed distinctive themes in the top versus average performers. Common qualities shared by top performers included a shared sense of purpose, a hands-on leadership style, accountability systems for quality and safety, a focus on results, and a culture of collaboration.

Conclusions

Distinctive leadership behaviors and organizational practices are associated with measurable differences in patient-level measures of quality and safety.

Acad Med. 2007; 82:1178–1186.

Pockets of clinical excellence may be found in most academic medical centers (AMCs), but the leaders of these institutions face the challenge of trying to achieve consistently high performance institution-wide, that is, across a wide variety of clinical services. The sciences of quality improvement and patient safety have focused largely on isolated clinical units or *microsystems*, demonstrating the

Dr. Keroack is vice president and director, Clinical Practice Advancement Center at University HealthSystem Consortium (UHC), Oak Brook, Illinois.

Ms. Youngberg is vice president for insurance, risk, quality, and legal services, UHC, Oak Brook, Illinois.

Ms. Cerese is director of quality, Northwestern Memorial Hospital, Chicago, Illinois. During this study, she served as senior director for clinical process improvement, UHC, Oak Brook, Illinois.

Ms. Krsek is director of operational benchmarking and nursing leadership, UHC, Oak Brook, Illinois.

Ms. Prellwitz is assistant director of information architecture, UHC, Oak Brook, Illinois.

Dr. Trevelyan is lecturer in management, Harvard School of Public Health, Boston, Massachusetts.

Correspondence should be addressed to Dr. Keroack, University HealthSystem Consortium, 2001 Spring Road, Suite 700, Oak Bridge, IL 60523; telephone: (630) 954-2436; fax: (630) 954-5879; e-mail: (keroack@uhc.edu).

effectiveness of unit-based leadership, targeted interventions, and ongoing measurement.¹ This approach has led to spotty gains, with high performing areas existing side by side with underperforming ones. A role for senior leaders in elevating performance across all units has been suggested,² but learning in this area has been impeded by the absence of a methodology for measuring quality and safety in an objective way across an entire medical center. Because data on institution-wide quality and safety are lacking, the leadership practices and organizational structures associated with broad-based excellence have been discussed largely on the basis of opinion or analogy to other industries.^{3,4}

In this study, we developed a composite index of quality and safety based on patient-level data that spanned a wide variety of clinical programs. We then used qualitative methods to identify the organizational factors that distinguished superior performers from average ones. The goal was to link objective data on performance of AMCs in quality and safety with organizational structures and practices in a way that would begin to define a basis in evidence for the

management of these complex institutions.

Method

The study was conducted in two phases. In the first, we used patient-level data to construct a composite scoring system for quality and safety that would apply to the majority of patients in AMCs. In the second, without knowledge of the composite scores by the investigating team, we conducted site visits and reviewed documents of six institutions in order to discover factors that distinguished superior performers from average ones. The study was designed as an exploratory investigation, with the goal of identifying those variables associated with high performance that merited future study. The overall study design, statistical methods, and approach to site visits were reviewed and approved by a steering committee composed of leaders of AMCs, including chief executive officers (CEOs), chief medical officers, chief nursing officers, chief financial officers, and directors of quality improvement, risk management, and patient safety.

Phase 1: Development of the performance measurement scoring system

We constructed a composite index of quality and safety derived entirely from patient-level data on outcomes and processes of care. We obtained our data from the University HealthSystem Consortium (UHC), a member-owned alliance of 97 university teaching hospitals whose clinical service chiefs are also chairs of academic clinical departments in US medical schools. Hospitals that share their data explicitly permit analyses of these data and comparisons with other centers. We used data from the 24-month period of 2003 to 2004 for the 79 AMCs for which we collect data. We used, as a guide, the six attributes of an ideal health care system articulated by the Institute of Medicine: safe, effective, timely, efficient, equitable, and patient-centered.⁵ Measures of patient-centeredness and outpatient care were not included because consensus measures in these areas were not available at the time of the study; therefore, the final composite index was composed of measures of patient safety, mortality, effectiveness, and equity.

The safety domain was composed of selected measures of preventable complications as determined by the Agency for Health Care Research and Quality (AHRQ). These patient safety indicators (PSIs) are derived from hospital discharge abstract information and include conditions such as complications of anesthesia, postoperative sepsis, and infections attributable to medical care.⁶ Within the safety domain, each measure was weighted equally.

The mortality domain was composed of measures of mortality rates for selected inpatient diagnoses as defined by AHRQ in its inpatient quality indicators (IQIs).⁷ These include conditions such as craniotomy, stroke, and myocardial infarction. Because these conditions applied to fewer than 20% of inpatients at all study sites, the IQIs were supplemented with measures of risk-adjusted mortality from UHC's databases. (UHC has developed a risk-adjustment methodology based on logistic regression that assigns expected mortality at the patient level within diagnosis-related groups [DRGs]. The model employs the severity of illness class assigned by the All Patient Refined

Diagnosis-Related Groups [APR-DRG] software as an independent variable.⁸ It also incorporates variables for age, gender, race, admission source, and comorbidities, the last as defined by the Agency for Health Care Research and Quality [AHRQ] using the methods developed by Erlichhauser.⁹) The 20 highest-mortality product areas (groupings of DRGs in areas such as bone marrow transplant, cardiac surgery, medical oncology, etc.) were selected, bringing the total of inpatient discharges analyzed for mortality to over 70% at all sites. Each IQI was given a weight of one in the mortality domain, and the UHC product line composite was weighted at 20.

For the area of effectiveness, we used measures of compliance with evidence-based practices as defined by The Joint Commission (TJC) in its hospital core measures. These are process measures endorsed by the National Quality Forum for the care of acute myocardial infarction, heart failure, and community-acquired pneumonia. We scored compliance in a bundled fashion, that is, each patient was required to receive all measures for which they were eligible. The Hospital Core Measures were supplemented by measures of surgical readmission within 14 days for the same condition or for a complication (excluding burns and trauma). Each of the four measures was weighted equally within the domain.

Finally, equity of care was evaluated using the Hospital Core Measures from the effectiveness domain analyzed for disparities of care across gender, race (white versus nonwhite) and socioeconomic status (Medicaid and self-pay versus all other payers). The number of statistically significant disparities for each pairing of disease and category (nine in all) was ascertained using Student's *t*, with *P* set at < .05.

To identify AMCs that clearly stood out from their peers rather than construct a scale that discriminated among AMCs at any point along the range, the data for each measure (29 in all) within the domains of safety, mortality, and effectiveness were mathematically transformed to fit a normal distribution; goodness of fit was confirmed using the Shapiro-Wilk test. A score of four was assigned for performance within one standard deviation (SD) of the mean,

with a point added or subtracted for each SD of variance. Measure scores ranged from one (least favorable) to seven. For the domain of equity, scores were assigned according to the number of significant disparities detected (seven for no disparities, four for one, and one for two or more). Data were unavailable for some institutions on TJC Hospital Core Measures; in these cases, we imputed scores of four for effectiveness (assuming average performance) and seven for equity (assuming no disparities of care). Low case volumes in IQI metrics (<25 cases) were also given an imputed score of four (assuming average mortality). None of the institutions selected for site visits had imputed values for any measure.

The domains of safety, mortality, and effectiveness were each weighted at 30% in the overall score; equity, an untested measure, was set at 10%. We found that giving a weight of over 50% to any given domain negated the contribution of the other domains to the overall relative rankings.

Composite scores for operational efficiency and financial performance, such as cost per case-mix-index-adjusted discharge and operating margin, were created as screens to exclude extremely high cost or inefficient sites from further study. These measures did not factor into the overall composite score, but institutions scoring in the bottom quartile for either of these areas were excluded from consideration for a site visit.

Phase 2: Selection of study sites and identification of characteristics of top performers

We completed the scoring and ranking of Phase 1 in 2005 and conducted a series of site visits in the same year. We selected three institutions from among the top five and three comparison sites from the middle of the distribution for the purpose of site visits. We made our selections in order to achieve a level of geographic distribution and to allow representation of AMCs affiliated both with public and private schools of medicine. The primary site visit team consisted of five of the six authors (excluding L.W.P.), individuals with expertise in the areas of executive leadership, medical leadership, nursing leadership, quality improvement, patient

safety, and risk management. We masked the performance scores of the six sites. Before each site visit, we requested a detailed list of documents that each institution then provided; these documents related to the structure of the quality program, job descriptions for key leaders, strategic goal setting activities, and board reports. During site visits, the team attempted to verify that what appeared on paper was evident in practice. In addition, these and other documents requested served as corroborative material for significant statements made during the interviews.

Site visits consisted of 1.5 days per visit gathering information via a wide range of interviews. The site-visit team interviewed the CEO, members of the governing board, the chief medical officer, the chief nursing officer, the chief

financial officer, clinical department chairs, division chiefs, nursing unit directors, medical unit directors, residents, and selected members of the medical and nursing staffs. The interview questions were designed not only to generate a comprehensive view of factors affecting performance in quality and safety but also to identify corroborating examples of specific behaviors (List 1). Asking the same questions, though not necessarily in the order listed, of all interviewees in a formal setting lent an element of structure to the interviews. Because we felt it was critical to discover what was important in the mind of the interviewee, we asked each of them a series of open ended questions in each topic area.

We identified the themes reported in this article through inductive analysis of the

interview notes and documentary materials using standard qualitative techniques. We patterned our qualitative research approach after the work of both Glaser and Strauss¹⁰ on grounded theory and Collins and Porras¹¹ on the paired approach to the study of organizational performance. The goal was to allow theories to emerge from the findings of the interviews rather than from any preconceptions or beliefs. We submitted our notes, which captured the behaviors, words, and ideas of those interviewed, to a contracted health care consultant who had not participated in the visits. He analyzed the notes, extracted common findings, and then tested these findings for consistency and reinforcement against the findings of both the interviewees of a given institution and the documents reviewed. The site visit team synthesized the findings into themes, which were subjected to iterative analysis and discussion until the team members reached a consensus. The team gave emphasis to observed behaviors rather than statements of general practice. Because some of the characteristics of top performers were occasionally present in the comparison sites, the team adopted a rule that a given finding would be considered noteworthy only if it were present in all three top performers and in no more than one of the comparison sites. The team reported the findings from each site visit as case reports. We reviewed these reports first with senior leaders at each site for confirmation of the accuracy of the facts on which we based our conclusions, and then with the steering committee involved in the design of the study, who confirmed their saliency.

Results

The individual measures used in the composite scoring system and the summary results across the 79 AMCs scored are shown in Table 1. Wide variations were seen across the measures of safety, mortality, effectiveness, and equity. The composite scores for all AMCs with 95% confidence intervals are shown in Figure 1. The distribution forms an S-shaped curve, with only a handful of institutions demonstrating scores distinctly higher or lower than their peers; the scores were placed into five groupings for purposes of reporting to the scored institutions. Each of the top five performers scored above the median

List 1

Areas of Leadership and Organization Explored During Interviews, Including Sample Questions, 2005

1. Executive leadership engagement around quality and safety

- In what ways are senior leaders involved in issues of quality and safety?

2. Strategic planning and goal setting

- How are goals developed for the institution?
- How are goals shared with departments and units?

3. Translating strategy into tactics

- How are strategic goals translated into practice and specific initiatives at the grassroots levels?

4. Systems of accountability for quality and safety goals

- Who is accountable for performance in quality and safety at the department and unit levels?
- Have any clinical faculty members had their privileges terminated in recent years because of quality issues?
- What information does the board receive regarding quality and safety?

5. Organizational expertise in performance improvement

- What is the role of the quality department in the institution?
- Describe a project in your institution to improve quality or safety that was successful.

6. Professionalism and the culture of quality and safety

- How would you describe the relationship between physicians and nurses in your institution?
- Provide an example where physicians and nurses worked together to solve a problem.

7. Use of information and technology

- What technologies has the institution adopted to advance quality and safety?
- How is information on quality and safety shared with the board, senior clinical leaders, and executive leaders?

8. Communication (internal and external)

- How do front-line employees learn about priorities for quality and safety?
- How do employees share their concerns about quality and safety issues?

9. Patient centeredness

- Are patients involved at all in standing committees?

10. General summary questions

- If we were to find out that this institution was a top performer, what factors do you think allowed it to achieve that level of performance?
- If we were to find out that this institution was average, what factors do you think held it back?

Table 1

Performance of 79 Academic Medical Centers on Individual Measures of Quality and Safety, as Components of a Composite Index, 2003–2004

Safety domain (30% weight): Agency for Healthcare Research and Quality (AHRQ) patient safety indicators (PSI)	Rate per 1,000 cases, average (range)	SD
Complications of anesthesia (PSI 1)	1.16 (0–29.71)	3.32
Death in low mortality DRGs (PSI 2)	0.87 (0.14–2.96)	0.52
Decubitus ulcer (PSI 3)	22.38 (9.47–51.27)	7.67
Failure to rescue (PSI 4)	124.43 (77.0–194.86)	24.16
Foreign body left in during procedure (PSI 5)	0.18 (0–0.75)	0.11
Iatrogenic pneumothorax (PSI 6)	1.51 (0.46–3.14)	0.58
Infection caused by medical care (PSI 7)	5.12 (2.33–11.86)	1.96
Postoperative hip fracture (PSI 8)	0.28 (0–1.05)	0.24
Postoperative hemorrhage or hematoma (PSI 9)	3.29 (0.72–6.56)	1.17
Postoperative physiologic or metabolic derangement (PSI 10)	2.87 (0–10.92)	2.02
Postoperative respiratory failure (PSI 11)	6.98 (1.73–25.94)	4.19
Postoperative pulmonary embolism or deep vein thrombosis (PSI 12)	16.02 (5.64–37.71)	5.61
Postoperative sepsis (PSI 13)	13.16 (2.93–45.30)	7.49
Postoperative wound dehiscence (PSI 14)	2.66 (0–7.72)	1.45
Accidental puncture (PSI 15)	6.19 (1.30–15.48)	3.34
Mortality domain (30% weight): AHRQ inpatient quality indicators (IQI)	Rate per 100 cases, average (range)	SD
Coronary artery bypass graft (IQI 12)	4.30 (0.52–11.81)	2.10
Craniotomy (IQI 13)	7.18 (0.66–23.33)	4.70
Acute myocardial infarction (IQI 15)	7.28 (2.74–16.16)	2.11
Congestive heart failure (IQI 16)	3.69 (0.71–7.46)	1.41
Stroke (IQI 17)	14.41 (5.98–27.97)	3.56
Gastrointestinal hemorrhage (IQI 18)	3.37 (0.90–6.39)	1.23
Hip fracture (IQI 19)	3.45 (0–10.26)	1.67
Pneumonia (IQI 20)	7.08 (2.96–11.89)	2.20
Acute myocardial infarction without transfers (IQI 32)	7.50 (2.74–16.03)	2.41
UHC product line composite*	1.00 (0.58–1.97) [†]	0.21
Effectiveness domain (30% weight): Joint Commission Hospital Core Measures	% Patients receiving all indicated measures, average (range)	SD
Acute myocardial infarction	76.09 (28.99–96.25)	13.18
Heart failure	42.35 (6.67–76.18)	18.52
Community-acquired pneumonia	57.56 (28.65–82.52)	11.73
Surgical readmission rate	2.88 (1.85–4.07) [‡]	0.53
Equity domain (10% weight): Disparities of care for disease and category	% Institutions with significant disparities [§]	
Acute myocardial infarction/gender	4.5	
Acute myocardial infarction/race	10.6	
Acute myocardial infarction/socioeconomic status	7.6	
Heart failure/gender	1.5	
Heart failure/race	13.8	
Heart failure/socioeconomic status	3.1	
Community-acquired pneumonia/gender	1.5	
Community-acquired pneumonia/race	13.8	
Community-acquired pneumonia/socioeconomic status	3.1	

* University HealthSystem Consortium (UHC) composite includes groupings of diagnosis-related groups in the following clinical categories: bone marrow transplant, cardiology, cardiothoracic surgery, gastroenterology, gynecology, heart transplant, kidney transplant, liver transplant, general medicine, general surgery, neurology, neurosurgery, orthopedics, pediatrics, pediatric surgery, surgical oncology, spinal surgery, urology, and vascular surgery.

[†] Observed/expected mortality ratio.

[‡] Percentage of cases readmitted within 14 days for same diagnosis or a complication of care.

[§] Student's *t* (two tailed, *P* < .05).

in each of the domains of safety, mortality, effectiveness, and equity. One of each of the comparison sites fell within groups two, three, and four in Figure 1.

The relative absence of discrimination among neighboring institutions in the middle part of the distribution is a reflection of the statistical approach of normalization of the data, which forces approximately two thirds of the scores for any measure into the average category. The failure of any institution to achieve a composite score greater than 70 is also a reflection of the same statistical method. Achieving a perfect score of seven in any given measure (>3 SD from the mean) would be expected to occur only 0.1% of the time; doing so in 29 separate measures would have a vanishingly small probability. The statistical approach did, however, achieve its goal of distinguishing a small number of institutions performing markedly better than the rest. Furthermore, the differences in scoring were associated with significant clinical differences. For example, a move from average to top levels of performance could result in 150 fewer deaths per year for a 500-bed AMC with a mix of cases typical to the cohort.

The findings from the six site visits were grouped into five themes that are summarized in List 2. For each theme, there was a group of findings related to factors found consistently in better performers, as well as a corresponding set of inhibitory factors, often present in the comparison institutions, that appeared to prevent them from achieving better performance.

Theme 1: A shared sense of purpose

The leaders of the top performers articulated the view that patient care was first among the missions of patient care, teaching, and research. This often took the form of sustained *patients first* campaigns or explicit values articulated in the mission and vision statements. Not only was there a greater level of focus on quality and safety by top performers in regular leadership meetings, reports, performance measures, etc., but leaders also routinely used leadership walk rounds or town meetings to address these issues. The leaders in the top performers also expressed a strong sense of dissatisfaction with the current state of quality and safety in their institutions. Instead of focusing on comparisons to

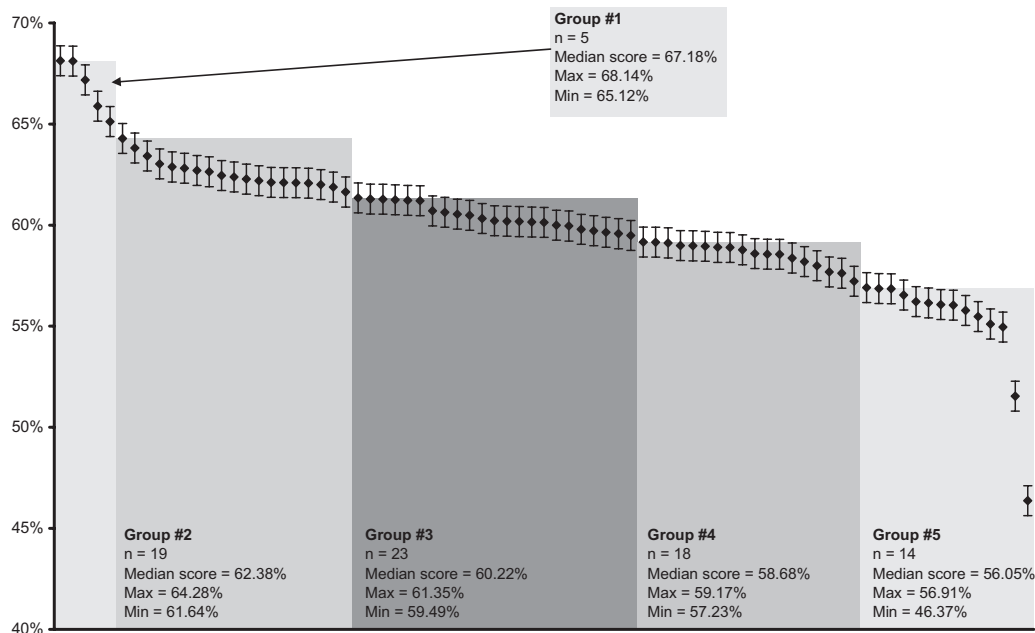


Figure 1 Composite scores for quality and safety among 79 academic medical centers, 2003–2004.

peers, they instead dwelt on gaps between the current state and a future ideal state toward which they were striving. Leaders charged with the missions of teaching and research either found this *patients first* point of view not to be in conflict with excellence in the other missions or expressed the opinion that the enterprise had achieved a balance among the missions.

Leaders in the top performers used a focus on service excellence to unify their institutions behind the *patients first* mission. The site visit team felt that this focus on service allowed for the engagement of both clinical and nonclinical staff in aspects of quality and patient safety. Rather than linking quality only with clinical outcomes, an area dominated by expert practitioners, the focus on service broadened participation in quality to areas such as timeliness, customer service, facility cleanliness, and safety. This enabled all staff members to be responsible for something which advanced the institutional mission. Lastly, the institution saw excellence in service, quality, and safety as a source of strategic advantage within a highly competitive marketplace, rather than simply “the right thing to do.” Two of the top three performers were in markets served by other AMCs, and all had experience with public reporting initiatives.

In contrast, the comparison institutions experienced unresolved conflicts among the missions of patient care, teaching, and research. In particular, the clinical department chairs were inconsistent at these sites in their interest in the patient care mission, and leaders were either unable or unwilling to address this inconsistency. Leaders at the comparison sites were typically satisfied with the current state of quality and safety, and they often pointed to external awards and other recognitions, in spite of signals from more detailed data sources that quality or safety problems persisted. For some, clinical quality was implied by specific scientific accomplishments, and measures suggesting average performance were treated with disbelief. Service excellence initiatives were isolated to individual clinical units and not used as an organizational imperative. Finally, leaders of the comparison sites tended to view quality and safety as part of professional ethics or “the right thing to do,” but not as a requirement for the strategic survival of the institution.

Theme 2: Leadership style

The CEO of each of the top-performing institutions was passionate about improvement in quality, safety, and service and had an authentic, hands-on style. All were routinely referred to by front-line staff by their first names, and staff were able to tell frequent stories

about personal encounters with the CEO that had reinforced the *patients first* vision. The CEOs themselves were skilled at storytelling, choosing to relate recent personal observations that either articulated the vision or highlighted gaps to be closed. They were frequent visitors to patient care areas, either as part of structured leadership walk rounds or as unscheduled observers. This personalized approach manifested itself in one case when the CEO chose to step in to park cars during a staffing shortage among the valet parking staff. Many of the staff cited this as an example of the CEO’s behavior in putting patients first.

In two of the top performers, the organizational structure minimized conflict among missions, in that both the hospital executive and clinical department leadership reported to a single CEO. In the third, the more traditional separation among components (hospital, school, and faculty practice plan) was managed collaboratively to reduce tensions among the components via a weekly operational meeting of the three leaders. All three top institutions were led as an alliance among the department chairs and executive leadership, with joint participation in strategy, program development, and performance improvement. The chief medical officer played a key role in working with the chairs to integrate their participation in system improvements.

List 2

Key Findings from Qualitative Analysis of Three Top-Performing Organizations**1. Shared sense of purpose:**

- Hospital leaders articulate that *patient care comes first*.
- Leaders are dissatisfied with the current state of quality and safety.
- Service excellence is added to the focus on quality and safety.
- Service, quality, and safety are seen as a source of competitive advantage.

2. Leadership style:

- The CEO is passionate about service, quality, and safety, and has an authentic, hands-on style.
- Everyday events are connected to the larger purpose through stories and rituals.
- Governance structures and practices minimize conflict between missions.
- The institution is led as an alliance between the executive leadership team and the clinical department chairs.

3. Accountability system for service, quality, and safety:

- Prioritizing, developing measures, and setting goals are centralized, and the tactics to improve are decentralized.
- The chairs accept responsibility for quality and safety within their departments.
- There is accountability, innovation, and redundancy at the unit level.

4. A focus on results:

- There is a relentless effort to improve, employing performance against external standards as a measure of success.
- Results outweigh the approach to performance improvement.
- There is a focus on human behavior and work redesign as the keys to improvement.
- Technology is employed as an accelerator and not as a substitute for work redesign.

5. Collaboration:

- Collaboration characterizes the relationships between administration, physicians, nurses, and other staff.
- Recognition of employee contributions at every level is frequent.
- Employees value each other's critical knowledge when problem solving.

Interestingly, top-performing institutions did not use financial incentives in the form of bonuses for high levels of performance to drive the quality agenda or to engage or reward key staff. Rather, in top-performing institutions, there seemed to be a strong but subtle pressure exerted on clinicians to conform to the values of the institution specific to quality, safety, and service. In one top-performing institution, the site visit team noted that the annual renewal of physician contracts was a time to address alignment with the institutional vision. In another (an institution with a mix of employed faculty and private practitioners), we were told that physicians who did not accept the institution's commitment to quality, safety, and service would see fewer referrals from colleagues.

The comparison institutions had less clarity at the leadership level, arising in part from tensions among the multiple missions. The CEOs were often unsure of their leadership roles in quality and safety

or had placed priority on other aspects of organizational performance, such as growth in market share, major building projects, or cost containment. Staff reported seeing them only rarely in patient care areas and did not feel free to raise quality or safety issues to them directly. Quality, safety, and service were discussed as abstract concepts and not regularly connected to personal stories of achievement or opportunity. Chairs often felt free to opt out of quality and safety initiatives, and the role of the chief medical officer was less prominent. A culture of "all persons for themselves" replaced the concept of alliance management in these institutions.

Theme 3: Accountability system for quality, safety, and service

The top-performing institutions demonstrated a blend of central control and decentralized responsibility. The development of strategic priorities and measures of success were developed by the central quality committee of the

institution, at which there was active board participation. On the other hand, tactics for improvement were typically entrusted to the unit leadership. In addition, dozens of grassroots projects emerged with little or no prodding from the central quality-oversight committee. These led to a much greater level of spontaneous innovation at the unit level, which occasionally was redundant and even unsophisticated. One obstetrical unit set as its goal to have no Apgar scores lower than five in full-term neonates without malformations and embarked on several initiatives with varying results before finally arriving at a set of practices that achieved sustained success. These local initiatives typically were celebrated and recognized by the leaders of the institution, even if they were not consistently supported by the central quality infrastructure.

In all top performers, the clinical department chairs accepted responsibility for quality, safety, and service in their departments, and this sense of accountability was transmitted to service chiefs and unit medical directors. In many cases, chairs insisted they be involved in such efforts and designed forums for problem solving and idea generation. Actual employment of the chairs and faculty by the institution was not necessary to achieve top performer status; one of the three top institutions used a mixed staff model, in which chairs and faculty could be either institutional employees or private practitioners.

The boards of top-performing institutions were more fully engaged in ensuring accountability for the quality, safety, and service agenda. Whereas board members from all six institutions expressed a commitment to quality and safety, those in the top performers were able to translate their values into effective oversight because of those institutions' commitments to transparency of information. Board members of top performers routinely saw data on quality and safety performance. In one case, select members of the board had access to the institution's scorecards via the Internet and received details about all sentinel events within 24 hours of their occurrence. In addition, there was evidence that the top-performing boards had insisted on follow-up and resolution of identified issues, often assigned as the responsibility of the affected department

chair or division chief, confirming that problems had been successfully resolved.

In the comparison sites, a much smaller number of local projects arose in individual units, and their champions typically felt unappreciated and disconnected from the larger strategy of the institution. Further, involvement by the chairs in the average-performing institutions was more inconsistent and not expected. Board members tended to see only summary reports from the quality department and were not expected to play a role in ensuring follow-up in areas of underperformance. Leaders chose to work only with cooperative chairs, while allowing the reluctant or unconcerned to avoid involvement. In the comparison sites, quality initiatives were both fewer and less widespread, and their execution was seen as the responsibility of the central quality committee. In its most extreme form, quality and safety were identified with a single individual or small group, absolving others of responsibility.

These comparison institutions also tended to gravitate to complex statistical methods, such as Six Sigma or Lean Toyota, perhaps in response to their difficulties in engaging the clinicians broadly. The result of committing to these complex methods was a small number of projects, rigorously executed, which seemed to dampen the initiative of less expert front-line staff. Although these sophisticated methods were familiar to quality leaders in the top performers, they were regarded as one of a number of possible tools to consider in problem solving.

Theme 4: Focusing on results

The top-performing institutions demonstrated a vast appetite for measurement across a wide variety of patient programs. There was an abiding sense that improvement was both necessary and possible, and these institutions routinely set goals using external standards to measure success. Measures were shared broadly with managers at all levels as well as the governing board in an easy to understand format; these often took the form of scorecards that were disseminated electronically. Clinicians often took pride in developing novel measures when there were not readily available national standards. Measurement was not an end

in itself, however. It supported the fundamental activity of work redesign to improve outcomes and reduce error. Although technology was often applied to accelerate gains, it was not used as a replacement for careful design of processes of care.

In comparison institutions, there was an emphasis on making gradual gains compared with past performance among those willing to collaborate with the quality and safety department, whereas areas of chronic underperformance often went unaddressed. The ability to come to consensus around optimal work redesign and system improvements often was lacking, allowing for wide discretion in the approaches to common clinical problems. There was also a sense that technology might substitute for attention to work redesign: quality and safety improvements were often a hoped for consequence, for example, of an electronic record or new building project, rather than flowing from a critical examination of current systems.

Theme 5: Collaboration

In top institutions, collaboration characterized the relationships among administration, physicians, nurses, and other professional staff. Multidisciplinary teams were the rule in efforts to improve; when solving problems, there was deference to expertise and situational knowledge rather than rank or position. These AMCs also seemed to be more regular in recognition of employee contributions at all levels. One had even set aside monies (\$170,000 in all) for small grants that were awarded to frontline staff to implement their best ideas to improve care.

Comparison sites were more likely to have formal initiatives in place to foster greater collaboration between physicians and nurses, perhaps recognizing this shortcoming in their cultures. Staff were more likely to feel that senior leaders did not work with them or understand their perspectives. In its worst manifestations, there were examples of rivalry and credit grabbing among disciplines, as well as a tendency of nurses to work around reluctant physicians rather than involving them.

Summary of findings and themes

The overall picture that emerged from these five themes was clear and

consistent. Top-performing institutions had settled on the concept of the primacy of patient care and were dissatisfied with the current state of quality, safety, and service. Their senior leaders reiterated this vision through stories, practices, and decisions that reinforced the core values. Clinical chairs and service chiefs accepted accountability for quality, safety, and service on their units and were willing to be measured against external standards. Innovation at the unit level was encouraged and celebrated. The institution was focused on defining new levels of excellence in the context of a culture that stressed mutual respect and professional responsibility.

Discussion

The current study used discharge abstract information to construct a composite index of quality and safety at an institutional level based solely on patient-level data. The scoring system effectively discriminated among 79 AMCs in ways that were clinically important. Qualitative analysis of data from the six site visits suggests that top performers in this system had distinctive leadership practices and organizational structures that were typically absent in average performers. Furthermore, using the scores to select a cohort of AMCs for qualitative study allowed the identification of those factors that were associated with superior performance, permitting a discussion of *evidence-based management* that is based on more than opinion.

Public reporting of quality and safety information has led to new levels of transparency for consumers and concern on the part of providers. This concern is at its height when composite measures are equated with quality at the institutional level in published rankings. One prominent scoring system relies heavily on case volume, adoption of specific technologies, and surveys of reputation; it has been criticized for correlating poorly with measures of evidence-based medicine (EBM).^{12,13} Another relies exclusively on risk-adjusted mortality for Medicare patients.¹⁴ A third scores institutions on compliance with evidence-based measures pertaining only to a small fraction of inpatient diagnoses.¹⁵

We suggest that the scoring system for this study is an improvement over

currently available institutional ratings of quality and safety. This system focused on outcomes at the patient level across a wide variety of patient types, using as a guide the six attributes of an ideal health care system articulated by the Institute of Medicine: safe, effective, timely, efficient, equitable, and patient centered.⁵ The data came from all payers and included measures of safety (derived from the AHRQ patient safety indicators), effectiveness (derived from TJC Hospital Core Measures performance and readmission rates), and risk-adjusted mortality (covering over 70% of inpatients). Because the risk-adjustment model for mortality corrected for variation based on race, gender, and socioeconomic status, disparities in care were assessed using process measures (disparities in performance around TJC Hospital Core Measures).

In spite of devising a composite scoring system more comprehensive and patient focused than others that are currently available, there remain several shortcomings to our approach. The AHRQ patient safety indicators have been tested for face and construct validity, but they have not been tested extensively for sensitivity and specificity, and thus they may be prone to underreporting bias or coding inaccuracies.⁶ The absence of standardized outcome measures in the ambulatory setting limited the focus of the study to inpatient care, leaving issues of ambulatory care quality and coordination largely unexamined. Measures of patient centeredness were not included in the composite score because consensus measures of patient satisfaction or other aspects of the patient experience of care were not available at the time of the study. The measure of equity of care was not independently validated and does not address disparities in access to care. Finally, although the weighting of the various domains was designed to confer balance, small changes in weighting, as well as the natural variation of the component measures, may affect the relative rankings in any composite scoring system.¹⁶

The scoring system was devised as a means to an end rather than an end in itself; it served as the basis for the comparative organizational studies that constituted the second phase of the study. Even before the findings were developed

from the interview notes, the site visit team spent little time in coming to consensus about the status of the institution as top or average performers in all six cases, indicating a surprising level of clarity in distinguishing the two groups based on the behaviors of the leaders and the institutional approaches to measurement, accountability, and improvement. The team's initial impressions were confirmed by the findings developed from the interview notes by an individual who did not participate in the visits. Had it been impossible to distinguish top-performing institutions from comparison sites, it could have been due to problems with either the scoring system (lacking validity as a measure of overall quality and safety) or the site visit methods (failing to detect the characteristics top performers had in common). Because there were consistent and distinctive organizational features for institutions with the best scores, the two phases of the study were mutually reinforcing.

The study did not aim to delve into the educational activities of the top and average performers. In both top and average performers, residents were involved occasionally, but not consistently, in quality and safety initiatives. However, the limited contacts with residents make residency education in quality and safety an area for future study.

There have been a limited number of studies seeking to correlate leadership behaviors and other organizational factors with outcomes of patient care. Two recent reviews noted common problems associated with narrow definitions of performance, as well as the variability in approaches to studying organizational factors.^{17,18} In two of the more rigorous studies, Holmboe et al¹⁹ and Bradley et al²⁰ characterized the attributes of physician leaders and senior managers in hospitals caring for patients with acute myocardial infarction. They linked superior performance with personal engagement by the chief executive, a focus on setting measurable goals, and practices to build a collaborative culture. Meyer et al²¹ studied four hospitals that had achieved superior performance in mortality and length of stay and identified four common factors: (1) CEO commitment, (2) an explicit mission statement focused

on quality, (3) a collaborative culture, and (4) efforts to get real-time information in the hands of caregivers. Shortell et al²² investigated data on 693 medical groups and concluded that regular involvement in quality-improvement initiatives and systematic measurement of patient satisfaction was correlated with superior scores on a composite of quality measures. The same investigators found an association between measures of team culture and intensive care unit performance as measured by length of stay, nurse turnover, and employee satisfaction.²³ Prybil²⁴ surveyed 20 hospitals and found correlations between risk-adjusted outcomes and (1) board involvement in quality-improvement activities, (2) senior management incentives for quality, and (3) close working relationships among the board and the medical staff.

The findings of the current study are consistent with the work noted above and extend the understanding of the relationship between institutional leadership and consistent excellence in patient care. Although comparison institutions often demonstrated superior performance in a single area that was associated with committed local leaders, consistent performance across the institution presented a challenge that required the involvement of top leaders. The use of a comparison group of average performers in this study permitted the identification of factors present exclusively or largely in top sites while also uncovering factors that inhibited achievement in the comparison institutions. Finally, the in-depth interviews conducted with dozens of senior leaders at each site allowed a more detailed understanding of beliefs, behaviors, and tactics than could be obtained by analyses of databases or surveys alone.

The role of leadership at the level of clinical microsystems in achieving high levels of performance has been stressed by Nelson et al.¹ A shared sense of purpose as well as the use of measurement and feedback have been found to be indispensable to those units that succeed in improving quality and safety. The role of leaders at the institutional level can be viewed as providing a common vision, fostering principles of measurement and feedback, promoting shared learning, and

celebrating success across the component microsystems, a model that is consistent with the pervasive grassroots efforts fostered by leaders in the top performers of this study.²⁵ In contrast, average performers relied on more centralized management of individual projects, an approach which may reflect the failure to engage opinion leaders at the unit level.

The second phase of the present study also had its weaknesses. The small sample size (six sites) can be regarded only as a beginning—an exploratory study to identify organizational factors deserving further study. Furthermore, analyzing the institutions at a single point in time makes inferences about causation impossible; the factors identified were merely associated with high performance. Studies of changing performance over time will be most valuable in identifying factors that lead to improvement. The use of data from 2003–2004 during the interviews conducted in 2005 may have led to recall bias among the interviewees. Lastly, the absence in the study group of the poorest performers in the scoring system may have led to our missing important inhibitory factors.

In spite of these limitations, however, the study advances the understanding of the role of leadership in achieving clinical excellence. At a time when gaps in the delivery of care are decried by policy researchers and the public alike, and when there are calls for evidence-based management to complement EBM,² it is important to define which leadership behaviors are effective in influencing clinical outcomes. Explicitly defining quality, safety, and service as top institutional priorities; demonstrating passion and engagement by the CEO; ensuring consistent accountability for the clinical service chiefs and innovation at the unit level; focusing on measurement and feedback; and fostering interdisciplinary collaboration all were found to characterize institutions that had distinguished themselves from their peers in broad-based measures of patient-care excellence. The findings of the study strongly suggest that actions on the part of the CEO, senior executive leaders, and department chairs have effects that are measurable at the bedside.

Acknowledgments

The authors acknowledge the important contributions of Rajneesh Behal, MD, MPH, in the development of the scoring methodology and of John M. Orlando, MBA, in extracting key organizational findings and themes from the interview notes.

References

- Nelson EC, Batalden PB, Huber TP, et al. Microsystems in health care: part 1. Learning from front-line units. *Jt Comm J Qual Saf.* 2002;28:472–493.
- Kovner AR, Rundall TG. Evidence-based management reconsidered. *Front Health Serv Manag.* 2006;22:3–22.
- Reinertsen L. A theory of leadership for the transformation for health care organizations. Available at: (www.ihl.org/NR/rdonlyres/F7CD1EBD-C2EF-48A1-8A1C-D1DE9C8F02AE/0/TransformationalLeadershipJanuary2004.doc). Accessed August 15, 2007.
- Ruchlin HS, Dubbs NL, Callahan MA. The role of leadership in instilling a culture of safety: lessons from the literature. *J Healthc Manag.* 2004;49:47–58.
- Committee on Quality of Health Care in America, Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century.* Washington, DC: National Academy Press; 2001.
- Agency for Healthcare Research and Quality. Guide to patient safety indicators. Available at: (www.qualityindicators.ahrq.gov/archives/psi/psi_guide_rev2.pdf). Accessed August 15, 2007.
- Agency for Health Research and Quality. Guide to inpatient quality indicators. Available at: (www.qualityindicators.ahrq.gov/archives/iqi/iqi_guide_rev4.pdf). Accessed August 15, 2007.
- Averill R, Goldfield N, Hughes J, et al. What are APR-DRGs? An introduction to severity of illness and risk of mortality adjustment methodology. Available at: (<http://multimedia.mmm.com/mws/mediawebserver.dyn?6666660Zjcf61Vs6EVs666ncICOrrrrQ>). Accessed August 15, 2007.
- Agency for Health Research and Quality. Healthcare Cost and Utilization Project, Comorbidity software, version 3.0. Available at: (www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp#archives). Accessed August 15, 2007.
- Glaser BG, Strauss AL. *The Discovery of Grounded Theory: Strategies for Qualitative Research.* New York, NY: Aldine De Gruyter; 1999.
- Collins JC, Porras JI. *Built to Last: Successful Habits of Visionary Companies.* New York, NY: Harper Business; 1997.
- McFarlane E, Murphy J, Olmsted MG, Drozd EM, Hill C. America's best hospitals 2007 methodology. Available at: (www.usnews.com/usnews/health/best-hospitals/methodology_report.pdf). Accessed August 15, 2007.
- Williams SC, Koss RG, Morton DJ, Loeb JM. Performance of top-ranked heart care hospitals on evidence-based process measures. *Circulation.* 2006;114:558–564.
- Health Grades, Inc. Hospital report card mortality and complication based outcomes 2007 methodology white paper. Available at: (<http://www.healthgrades.com/media/DMS/pdf/HospitalReportCardsMortalityComplications2007.pdf>). Accessed August 15, 2007.
- United States Department of Health and Human Services. Hospital Compare—a quality tool for adults, including people with Medicare. Available at: (<http://www.hospitalcompare.hhs.gov/Hospital/Static/Data-Consumers.asp?dest=NAV|Home|DataDetails|ProfessionalInfo#TabTop>). Accessed August 15, 2007.
- Jacobs R, Goddard M, Smith PC. How robust are hospital ranks based on composite performance measures? *Med Care.* 2005;43:1177–1184.
- Scott T, Mannion R, Marshall M, Davies H. Does organizational culture influence health care performance? A review of the literature. *J Health Serv Res Policy.* 2003;8:105–117.
- Hoff T, Jameson L, Hannan E, Flink E. A review of the literature examining linkages between organizational factors, medical errors, and patient safety. *Med Care Res Rev.* 2004;61:3–37.
- Holmboe ES, Bradley EH, Mattern JA, Roumanis SA, Radford MJ, Krumholz HM. Characteristics of physician leaders working to improve the quality of care in acute myocardial infarction. *Jt Comm J Qual Saf.* 2003;29:289–296.
- Bradley EH, Holmboe ES, Mattern JA, Roumanis SA, Radford MJ, Krumholz HM. The roles of senior management in quality improvement efforts: what are the key components? *J Healthc Manag.* 2003;48:15–28.
- Meyer JA, Silow-Carroll S, Kutyla T, Stepnick LS, Rybowski LS. Hospital quality: ingredients for success—overview and lessons learned. Available at: (www.cmfw.org/usr_doc/761_Meyer_overview.pdf). Accessed August 15, 2007.
- Shortell SM, Schmittiel J, Wang MC, et al. An empirical assessment of high-performing medical groups: results from a national study. *Med Care Res Rev.* 2005;62:407–434.
- Shortell SM, Zimmerman JE, Rousseau DM, et al. The performance of intensive care units: does good management make a difference? *Med Care.* 1994;32:508–525.
- Prybil LD. Size, culture, and composition of high-performing hospital boards. *Am J Med Qual.* 2006;21:224–229.
- Kosnik LK, Espinosa JA. Microsystems in health care: part 7. The microsystem as a platform for merging strategic planning and operations. *Jt Comm J Qual Saf.* 2003;29:452–459.