Safety is part of quality: a proposal for a continuum in performance measurement

Vahé A. Kazandjian PhD MPH,1 Karol G. Wicker MHS,2 Nikolas Matthes MD, PhD, MPH, MSc3 and Sam Ogunbo PhD4

1President, Center for Performance Sciences, Elkridge, MD, and Professor (Adjunct), Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
2Director of Product Management, Center for Performance Sciences, Elkridge, MD, USA
3Senior Director of Research, Maryland Quality Indicator Project®, Elkridge, MD and Assistant Professor (Adjunct), Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
4Statistician, Maryland Quality Indicator Project®, Elkridge, MD, USA

Keywords
patient-level data, Quality Indicator Project®, safety of care

Correspondence
Dr Vahé A. Kazandjian
President
Center for Performance Sciences
6820 Deerpath Road
Elkridge
MD 21075
USA
E-mail: vkazandjian@mhaonline.org

Accepted for publication: 26 October 2007
doi:10.1111/j.1365-2753.2008.00960.x

Abstract

Objectives Safer care is a strategic priority for health care organizations worldwide. Yet, the measurement and evaluation of key processes and outcomes associated with safer care remains challenging, even with existing performance measurement indicators. The multinational Quality Indicator Project (QI Project®) data are analysed to [1] document the patterns of safety indicators used between 1999 and 2006 among hospitals in Asia, Europe and the USA; and [2] to identify trends in using both organization-level and patient-level data in hospital performance improvement.

Design and setting Retrospective data are used to ascertain how the use of safety indicators has changed in comparison to other QI Project® indicators. ‘ Continent’ rather than ‘hospital’ is used as the unit of analysis and P-values of the differences in use percentages across Asia, Europe and the USA are calculated.

Results There was a significant increase in the use of QI Project® indicators in Asia between 1999 and 2006. Measured as the mean percentage of usage, the safety versus ‘all other’ indicators’ increase in Asia was 43.7% versus 27% (P < 0.05) and 37.2% versus 24.4% (P < 0.05), respectively, during the study’s time period. The European participants used both safety and all other indicators less frequently, 14.7% versus 18% (P < 0.05) and 9.5% versus 19.8% (P < 0.05), respectively. Finally, USA hospitals demonstrated a larger difference in the decrease of QI Project® indicator use than European hospitals between the ‘safety’ and ‘all other’ indicators, 12.7% decrease for safety indicators and 7.1% for all others (P < 0.05). These findings are consistent with trends reported in a previous study.

Conclusion Traditional performance measures continue to assist hospitals in identifying crucial aspects of safety in the delivery of care. Building on the findings of a previous study, there are emerging trends in the type of measures used in hospitals in Asia, Europe and the USA pursuing the improvement of overall performance. The increasing use of patient-level data specifically, in tandem with organizational level indicators, may signal the continuum of measurement strategies, now still predominately in the USA but anticipated to be adopted both in Europe and Asia.

Introduction

Safety continues to constitute a leading dimension of accountability for health care organizations: be it the safety of patients’ care, safe environment for family members and visitors, or the safety of health care providers themselves – safety is the term most often used to describe activities hospitals and other health care organizations engage in to demonstrate their priorities for the provision of patient care.

Worldwide the interest in measuring adverse events and understanding their relationship to causing harm can be traced back to the time when systematic performance measures, often called ‘indicators of quality’, were first developed in the 1980s. While these initial indicators of quality were eventually superseded, at least in terminology, by what we now commonly refer to as ‘performance measures’ [1,2], the general philosophy remains: processes of care should be well-defined and mapped to allow for the identification of bottlenecks affecting processes’ efficiency,
Effectiveness of desired outcomes of care, or, in the case of safety, the very relationship between processes and actual or potential unsafe provision of care.

Revisiting and expanding the findings of a previous study of the utilization of patient safety indicators in the Quality Indicator Project® [3,4], this study further explores indicator utilization patterns of hospitals in the USA, Europe and Asia over a 7-year period from 1999 to 2006. The main objective is to analyse the patterns of indicator utilization for both quality and safety indicators regarding the contribution of aggregate level versus patient-, procedure-, or disease-level measures in changing environments across continents.

Are indicators of quality different than safety indicators?

Ever since the publication of the Institute of Medicine report in 1999 a scholarly debate about a possible distinction in the measurement of ‘quality’ and ‘safety’ in the provision of care has emerged. Similar to the distinction between quality indicators and performance measures often made in the 1990s, the distinction now seems to be revisited this time in the terms of quality and safety. While Donabedian’s initial conceptualization of quality as ‘appropriateness’ had quite deliberately incorporated ‘safety’ [5], the following transformation of his conceptual model into the design and practical implementation of performance measures has, perhaps unconsciously, created a distinction between quality and safety.

Quality, measured primarily as performance, has often focused either on the absence or presence of best practice. Consequently, most of the quality measures were designed as process measures to assess whether best practice was adhered to [6,7]. Safety, in turn, often triggered by catastrophic adverse events involving patient harm, initiated an investigation into the ‘root causes’. While initially focusing on assigning blame to explain and punishment to remedy harmful and untoward outcomes [8–10], the field has fully recognized that it is the environment and conditions within which a process (e.g. provision of health care) occurs that correlate more strongly with the goodness [5] of those processes (or system) rather than the people. Hence the term ‘system approach’ [11,12]. Today, as the result of the evolution of our understanding of the interrelationship between quality and safety, not only does safety constitute an integral part of quality but the very measures capturing various aspects of quality and safety are often the same. What differs, however, is how those measures are interpreted and what strategies are adopted to either improve the efficiency and effectiveness of the processes or eliminate adverse events.

While the QI Project® represents only US participants and the International QI Project only international participants, for the purposes of this paper the term QI Project® encompasses both US and non-US hospitals [13,14].

Methods and analysis

The sample and indicators

A total of 582 European, Asian and US hospitals participating in the QI Project® from 1999 to 2006 constitute the cohort for this study: 79 in Asia, 149 in Europe and 354 in the USA.

The utilization of QI Project® indicators by continent for both quality and safety is shown in Table 1.

Analysis

Data were analysed using SAS software version 9.1 for Windows (2006). Frequency counts of use for all indicators (including safety and all other) were generated for 1999 and 2006 for Asia, Europe and the USA. Student’s t-tests were used to assess the significance of the difference in use between 1999 and 2006 for both ‘safety’ and ‘all other’ indicators separately for the three continents. Significance levels were set at $P < 0.05$. The mean difference for indicator use between 1999 and 2006 were computed for both ‘safety’ and ‘all other’ indicators (Table 2).

Findings

The overall frequency of utilization of the QI Project® indicators has either remained constant or changed little among European and Asian participants between 1999 and 2006. In the USA, hospitals have used these indicators as ‘safety’ indicators less frequently than overseas participants, perhaps due to increasing external requirements for US hospitals to collect and submit safety data (Fig. 1) [15].

The results of this study show there was a significant increase in indicator use in Asia between 1999 and 2006. The average change in use in Asia was 16.7% for ‘safety’ indicators and 12.8% for ‘all other’ indicators. In contrast, US hospitals had a 12.7% decrease in the use of ‘safety’ indicators ($P \leq 0.05$). Similarly, European hospitals had a 3.3% and 10.3% significant decrease ($P \leq 0.05$) in the use of ‘safety’ and ‘all other’ indicators, respectively (Fig. 2).

With the QI Project’s® voluntary nature, it is conceivable that participation wanes especially when not every hospital can muster the resources to participate both in mandated and voluntary initiatives. The corollary of that interpretation is that those hospitals continuing to participate in voluntary initiatives constitute a self-selected cohort of hospitals that benefited either from the participation in multiple initiatives [16] or from their continuous participation in the QI Project® [17].

Table 1, however, highlights a number of indicators that continue to increase in popularity among all participants. Indeed, restraints, falls and resulting injuries are systematically and continuously monitored among QI Project® participants.

Discussion

While the number of participating hospitals in Asia and Europe increased between 1999 and 2002 and has remained relatively stable since, the number of participating hospitals in the USA has decreased from 1031 in 1999 to 354 in 2006. The interpretation of this dramatic decline in the US has to be placed in the context of a rapidly external regulatory environment. Beginning in 2002 the Joint Commission required hospitals to submit data on patient-level, disease-specific indicators. The Centers for Medicare and Medicaid Services (CMS) followed suit in 2004 with the Starter Set requiring hospitals to submit data or forfeit a percentage of their overall Medicare reimbursements [18–20]. A uniform measure set for both purposes has resulted – now commonly called the National Hospital Quality measures (NHQ) [19,21–24].
In 1998, the QI Project, the largest indicator-based performance measurement system for health care organizations, was one of the first to be certified as an ORYX non-core vendor and continues to serve more than 650 US hospitals [25–29] with their proprietary indicators analyzed for this study as well as the NHQ measures [30]. The QI Project developed a web-based software tool to collect data, provide hospitals with trend and comparative analysis and, following extensive data quality checks, submit the data on the hospitals’ behalf to both the Joint Commission for accreditation and CMS for full reimbursement. During the 4-year period between 2002 and 2006, many of the hospitals that had been participating in the QI Project for a decade or more shifted participation from the QI Project’s proprietary indicators to the NHQ measures. Nevertheless, many US hospitals continue to participate in the proprietary measure, which may well enjoy renewed attention should CMS move towards a reimbursement strategy to no longer cover hospital acquired conditions fully [31]. This would hasten hospitals to focus their measurement efforts in a much broader sense to capture adverse events more effectively rather than focusing on the narrowly defined populations of the NHQ measures.

This trend may well forecast what may occur elsewhere in the foreseeable future: an evolution in investigative capabilities and focus towards a mix of both aggregate and disease- and patient-level measurement. Such a mix would complete the investigative model of understanding both the processes themselves as well as the associated better practices. The feedback loop shown in Fig. 3 constitutes an important part of such an evolution, such as hospital Boards or Ministry of Health during strategic planning. The next section further discusses patient- and disease-level measures.

Table 1: Frequency of QI Project® sub-indicator use as ‘safety’ indicators in Asia, Europe and the USA

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Asia (%)</th>
<th>Europe (%)</th>
<th>USA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibiotic prophylaxis for hip arthroplasty within 2 hours prior to incision</td>
<td>50.00</td>
<td>49.63</td>
<td>60.76</td>
</tr>
<tr>
<td>Antibiotic prophylaxis for vaginal hysterectomy</td>
<td>5.62</td>
<td>5.36</td>
<td>6.91</td>
</tr>
<tr>
<td>Unscheduled acute care re-admissions within 15 days for the same or related condition</td>
<td>45.67</td>
<td>47.67</td>
<td>56.70</td>
</tr>
<tr>
<td>Unscheduled acute care re-admissions within 31 days for DRG 127, heart failure and shock, or related condition</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Unscheduled admissions following ambulatory cardiac catheterization</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Unscheduled admissions following all other ambulatory operative procedures</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Unscheduled returns to intensive care units</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Unscheduled returns to the operating room</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Acute care physical restraint events</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Acute care physical restraint events due to risk of falling</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Acute care inpatients with one or more physical restraint events</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Documented falls in acute care</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Documented falls in acute care resulting in injury with Severity Score 1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Unscheduled returns to the ED within 24 hours</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Unscheduled returns to the ED within 24 hours, resulting in an inpatient admission</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Unscheduled returns to the ED within 72 hours</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Unscheduled returns to the ED within 72 hours, resulting in an inpatient admission</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

*‘Safety’ indicators constitute a subset of the total 715 measures used in the QI Project®.

Table 2: The mean percentage of usage and safety indicators by country

<table>
<thead>
<tr>
<th>Sponsor</th>
<th>Safety**</th>
<th>Mean percentage of usage (SD)</th>
<th>Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asia</td>
<td>0</td>
<td>27.38 (22.9)</td>
<td>10.3*</td>
</tr>
<tr>
<td>Europe</td>
<td>0</td>
<td>14.72 (21.9)</td>
<td>-12.7</td>
</tr>
<tr>
<td>USA</td>
<td>1</td>
<td>20.70 (16.6)</td>
<td>-10.3*</td>
</tr>
</tbody>
</table>

*Values are significant at the 0.05 level.

**0 = all other indicators; 1 = safety indicators.
Indicators of quality assess magnitude (events, frequency of processes, etc.). Through both statistical and clinical decision-making processes, changes in the magnitudes of measurement over time assist organizations in identifying priorities for improvement. For that reason alone, comparative analysis remains essential, be it to an organization’s past performance or the performance of peers (while adjusting for confounding variables, if necessary).

In the case of safety indicators, however, the philosophy appears entirely different: adverse events, often described with terms ranging from ‘no events’ to ‘near misses’, may not require comparative data. Indeed, it could even be proposed that for some safety measures one event is too many. Risk management and risk managers are primarily focused on those singular outcomes. For example, while it was not necessary to establish how many wrong doses of chemotherapy drugs were administered to a patient who developed kidney failure, it was sufficient to know that one patient had developed kidney failure because of wrong chemotherapy dosage. It is the very nature of safety measure events to occur with low frequency, although the associated outcomes can be catastrophic.

As the scientific literature has focused increasingly on the importance of near-misses, even the potential for errors, a basic reconsideration of the initial distinction between ‘quality’ and ‘safety’ indicators seems in order. Seminal works on errors resulting from the provision of a service in any industry, have well established that errors can occur during any process. Therefore, it appears of much greater importance to understand the environment, structures, processes, as well as the attitudes of the people themselves rather than the outcomes defined as either quantifiable or qualifiable events. This accounts for the rapprochement between the concepts on the
one hand and the mechanics of defining and designing quality indicators on the other. When analysis of a process is required to understand whether best knowledge at the time (evidence-based practice) was followed or whether the process suffered from inherent predispositions to undesirable outcomes (such as errors), the very distinction between ‘quality’ on the one hand and ‘safety’ indicators on the other becomes noticeably blurred.

Patient/disease level analysis

The use of patient-level and disease-specific performance and patient safety measures has risen enormously in the US over the past few years. This trend has been accelerated by renewed attention to external accountability. Both the Joint Commission and the CMS require hospitals to submit patient-level data on clinical conditions including heart failure, acute myocardial infarction, pneumonia, surgical infection prevention, pregnancy and childhood asthma. Measures of outpatient and psychiatric care are forthcoming.

Although the use of these measures has been driven by regulatory requirements and the need for external accountability (i.e. holding data to precise and, if necessary, risk-adjusted, interfacility comparisons for public release and performance-based reimbursement), the measures are also beneficial to the hospitals themselves. Abstracting and reporting on patient-level measures allows for real-time data analysis while aggregate measures do not. Take, for example, a hospital that has identified poor timing of antibiotic prophylaxis prior to surgery. (Table 1 shows antibiotic prophylaxis as one of the safety measures used by hospitals in this study.) If the data are aggregated, it is not possible to identify the patients who had not received antibiotics in time and to understand the reasons why. This can only be achieved using patient-level data, where one can establish whether the type of surgery, patient characteristics, co-morbidities, operating surgeon, weekend or night-time emergency surgery, or other factors are associated with poor timing of antibiotic prophylaxis.

Aggregate measures do not allow for real-time analysis. Ultimately, patient-level and detailed data must be collected and analysed in order to pinpoint the reasons for variation in performance. Especially with patient safety issues associated with adverse outcomes, instant detailed data analyses offer great benefit for more timely and accurate understanding of the variation in performance.

Recommendations

The objective of this paper was to show that existing performance measures, using the QI Project® as an example, can and are being used worldwide for the quantification of safety in hospitals. While it is widely recognized that organizational readiness, technology [32,33] and local accountability mandate [34] influence the types of measures used as well as their frequency and sustainability, our findings strongly suggest that, strategically, a parallel measurement system may not be needed for ‘quality’ and ‘safety’ assessment and evaluation. On the contrary, a continuum is proposed: hospitals and other health care organizations should expand and supplement the types of measurement tools to address specific areas prone to errors or system failure.

An extrapolation of this recommendation addresses the extension of existing performance systems and strategies to patient- and disease-level indicators. While aggregate level analysis provides indispensable trend and pattern profiles for organizations, patient-level data support analyses into causes, especially when infrequent and unacceptable errors occur. Recent developments, including CMS ‘never events’, seem to pave the way for such a continuum of aggregate level and patient-level measurements.

In this study, and perhaps for the first time in the literature, a multi-continental trend is identified in the use of indicators as a reflection of improved investigative knowledge and evidence-based practices. The trends can be perceived along a continuum ranging from aggregate, organizational level indicators as
measures of the appropriateness of care, to the increasing adoption of patient or disease-level measures to understand both the overall quality and the safety of care.

In conclusion, existing performance measurement systems assist health care organizations in addressing both their overall performance and specific key aspects of safety of care. As these organizations continue in their use of aggregate level data, patient and disease-level measures provide the next level of detail and specificity in understanding what happened and why it happened. Until then, hospitals worldwide should explore how existing performance measurement systems can help them address the appropriateness of the care they provide by combining quality and safety [35].

Acknowledgements

The authors wish to thank Dana Bonistalli, Executive Assistant, Center for Performance Sciences, for her assistance in the preparation of this article.

References