Building a Culture of Safety and Quality: The Paradox of Measurement

Katherine Bliss

The use of performance measures and quality metrics in health care has become increasingly common over the past 30 years (MacLean et al., 2018). These measures are integrated into almost every healthcare setting and include measures such as mortality rates for specific conditions or after surgeries, hospital readmission rates, patient experience, trends over time, cost, and more (Khera & Krumholz, 2018; Medicare.gov, n.d; Toevs et al., 2017). Pay for performance, which financially incentivizes healthcare providers to achieve better patient outcomes, and public reporting, which provides information to patients regarding the quality of providers and organizations, have been major components of performance and quality measurement in health care (Berenson & Rice, 2015; MacLean et al., 2018). Moreover, the bipartisan support of the transition from fee-for-service volume-based care to value-based care ensures that measurement of cost and outcomes is likely to accelerate.

Linking reimbursement to measurement creates substantial incentives for organizations to perform well on quality and cost metrics. Hospital Value-Based Purchasing (HVBP), for example, extracts 2% from Prospective Payment System hospitals’ reimbursement (those hospitals reimbursed by diagnosis-related groups) with the opportunity to recoup some, all, or more of that cut based on how the hospital performs on 20 measures. While the goal of such measurement-referenced initiatives is to drive clinical behavior to high quality, it is unclear whether the use of performance and quality measurement consistently leads to better patient outcomes (Eijkenaar et al., 2013; Himmelstein et al., 2014; Powell et al., 2011). Nurse managers and leaders work in the complex interface among these forces and requirements.

Specifically, organizational culture influences healthcare providers’ clinical decision-making (Swensen & Monta, 2019), and nurse leaders both shape and are shaped by that culture. This impact of culture includes implicit or explicit pressure to achieve performance standards, particularly those that are publicly reported or linked to payment. Optimally, this pressure impels actions that yield positive patient outcomes. This, however, is not ensured. In some cases healthcare providers have delayed

Megan Chambers

Betty Rambur

Performance measures and quality metrics are widely used in health care. Yet, evidence has shown conflicting results as to whether this type of measurement consistently leads to improvements in quality of care and patient outcomes. The aim of this study was to examine U.S. emergency and intensive care nurses’ experiences with measurement-driven clinical behavior, with a focus on measurement-driven harm. Qualitative analysis revealed three major interacting categories of negative, unintended consequences of performance measurement: overtreatment and unnecessary care, undertreatment, and inefficiency. The resulting framework is offered to support nursing leaders’ efforts to enhance safety and quality, minimize measurement-related harm and inefficiencies, and contribute to the positive evolution of performance measures in health care.

Building a Culture of Safety and Quality: The Paradox of Measurement

Katherine Bliss

The use of performance measures and quality metrics in health care has become increasingly common over the past 30 years (MacLean et al., 2018). These measures are integrated into almost every healthcare setting and include measures such as mortality rates for specific conditions or after surgeries, hospital readmission rates, patient experience, trends over time, cost, and more (Khera & Krumholz, 2018; Medicare.gov, n.d; Toevs et al., 2017). Pay for performance, which financially incentivizes healthcare providers to achieve better patient outcomes, and public reporting, which provides information to patients regarding the quality of providers and organizations, have been major components of performance and quality measurement in health care (Berenson & Rice, 2015; MacLean et al., 2018). Moreover, the bipartisan support of the transition from fee-for-service volume-based care to value-based care ensures that measurement of cost and outcomes is likely to accelerate.

Linking reimbursement to measurement creates substantial incentives for organizations to perform well on quality and cost metrics. Hospital Value-Based Purchasing (HVBP), for example, extracts 2% from Prospective Payment System hospitals’ reimbursement (those hospitals reimbursed by diagnosis-related groups) with the opportunity to recoup some, all, or more of that cut based on how the hospital performs on 20 measures. While the goal of such measurement-referenced initiatives is to drive clinical behavior to high quality, it is unclear whether the use of performance and quality measurement consistently leads to better patient outcomes (Eijkenaar et al., 2013; Himmelstein et al., 2014; Powell et al., 2011). Nurse managers and leaders work in the complex interface among these forces and requirements.

Specifically, organizational culture influences healthcare providers’ clinical decision-making (Swensen & Monta, 2019), and nurse leaders both shape and are shaped by that culture. This impact of culture includes implicit or explicit pressure to achieve performance standards, particularly those that are publicly reported or linked to payment. Optimally, this pressure impels actions that yield positive patient outcomes. This, however, is not ensured. In some cases healthcare providers have delayed
necessary hospital readmissions within 30 days of discharge from the hospital (Fonarow, 2018), denied patients surgical procedures due to the risk of mortality (Toevs et al., 2017; Wadhera et al., 2017), and prolonged the life of critically ill, post-surgical patients until they were past the measurement period of 30 days postoperative (Rambur et al., 2013).

**Context**

The movement toward more intensive and extensive performance measurement was catalyzed with the landmark *To Err is Human* report, which shed light on the widespread safety issues in the U.S. health system. Yet, evidence suggests that 20 years after the report, little progress has been made (Castellucci, 2019a), and there is at least some evidence that metrics deemed as successful (e.g., lower hospital readmission rates) may have resulted in unintended deaths (Fonarow, 2018). Such unintended consequences are the basis of exploration in this study. Specifically, despite the near-universal nature of performance measures in the acute care setting, few studies explore nurses’ awareness of and experiences with these measures. This exploratory study was designed to address this gap. The aim was to understand better emergency department (ED) and intensive care unit (ICU) nurses’ awareness of performance measurement in their practice setting. A second aim was to illustrate their experiences of measurement-driven behavior, with a focus on measurement-driven harm, as the first step toward redesigns that can enhance the value of nursing, performance measurement, and performance-driven care.

**Methods**

A qualitative descriptive design (Sandelowski, 2000, 2010) using anonymous web-based data collection was used. Following a literature review, open-ended exploratory queries were developed. After obtaining approval from the local Institutional Review Board, snowball sampling, as well as advertisements on Facebook and the Arthur L. Davis nursing resource website, was employed for participant recruitment, a process that is detailed elsewhere (Chambers et al., 2020). Participant inclusion criteria included nurses working in the United States, age 21 or older. The study targeted ED and ICU nurses, but nurses in any specialty could choose to participate. Consistent with the standard approach in exploratory, hypothesis-generating, qualitative research, small sample size was intended. Before data collection, voluntary, informed consent was given by each nurse interested in participating in the study. Data were collected from June 18, 2019 to July 29, 2019 using Research Electronic Data Capture (REDCap) as the online interface and data collection entry and retrieval portal. Participants answered an interview guide containing four open-ended questions regarding performance measurement and eight demographic questions (see Table 1). After data collection, qualitative content analysis, the preferred analysis method for studies using a qualitative descriptive design (Sandelowski, 2010), was used. Since REDCap gathers data precisely as entered by the participants, all interview data were extracted as entered, then categorized into codes.

Next, themes were inductively developed and patterns analyzed. Study rigor followed the orientation of Lincoln and Guba (1985) through multiple steps. These included supporting credibility through transcription rigor, the search for confirming and disconfirming evidence, investigator triangulation, and support of dependability through a careful decision trail.

**Results**

All of the 17 study participants reported awareness of performance measurements within their healthcare setting. The most frequently reported measures were related to rapid treatment for selected conditions such as time of ED arrival to computed tomography (CT) scan, tissue plasminogen activator (TPA) administration or thrombectomy, or treatment with antibiotics. Other commonly recognized metrics included: (a) ED wait time, (b) time from patient arrival to percutaneous coronary intervention for patients with ST-elevation myocardial infarctions, (c) catheter-associated urinary tract infection (CAUTI) rates, and (d) central line-associated bloodstream infection (CLABSI) rates. No respondent cited any measures related to cost or HVBP purchasing measures in the Person and Community Engagement domain other than “patient satisfaction.”
Nine of the 17 study participants reported witnessing positive effects or patient outcomes related to performance and quality measurement. The metrics most frequently reported to have a positive influence were those pertaining to rapid treatment as well as perceived reductions in patients’ CAUTI rates, CLABSI rates, medication errors, ED wait times, and pressure ulcer rates.

All but two of the participants shared experiences of unintended, negative consequences of performance measurement or metric-driven harm they had witnessed within the healthcare setting. Content analysis of these data, followed by development of themes, fostered creation of a three-component framework by which nurse leaders can conceptualize interactions that fuel measurement-driven harm. These three components are overtreatment and unnecessary care, undertreatment, and inefficiency (see Figure 1). To illustrate, nurses’ time and attention were diverted to measured elements of care. Paradoxically, this could create overtreatment directed toward activities that enhance those measures, undertreatment in those that do not, and inefficient delivery of nursing care. Each component of the framework will be detailed, followed by practical implications for nurse leaders.

**Overtreatment and Unnecessary Care**

Nearly all nurses participating in the study reported seeing overtreatment or unnecessary care related to performance measures within their healthcare setting. Several participants reported witnessing harm associated with the same measures they had first reported as improving patient outcomes. For example, participants reported an increase in patients who received brain CT scans as well as TPA when stroke quality metrics were closely monitored. Reported harm included physicians who administered TPA to patients who did not have a stroke when stroke quality metrics, particularly time to treatment, were monitored closely. Unintended consequences included unnecessary radiation exposure and increased risk of bleeding. Unnecessary care also stemmed from routine standing orders that were not individualized to reflect the patient’s actual condition (e.g., a routine brain CT every 6 hours, even if the patient had unvaried mental status and care changes based on the CT results were not intended).

Routine chest x-rays in patients without any change in respiratory status and expensive workups that include routine pregnancy tests on postmenopausal women, routine electrocardiograms in stable hospitalized patients on a telemetry monitor, and treating patients with opioid medications for minor injuries that could be appropriately treated with anti-inflammatories were also described (the latter to maximize patient satisfaction scores). Participant responses also illustrate the potential for hospital policies to exacerbate the very problems they are trying to prevent (e.g., a facility’s requirement that nurses straight catheterize patients three times before inserting a Foley catheter for urinary retention, leading the study participant to question whether the policy helped improve or worsen CAUTI risk).

The following statement by an ED nurse illustrates respondents’ experience of the

---

**Table 1. Interview Guide**

1. Please describe patient care measures and measurement requirements you have seen in your practice setting.
2. Have these measures influenced clinical decision-making and patient care? Please describe any positive, neutral, or negative consequences of measurement-driven clinical behavior.
3. Cases have been described in which measurement-driven behavior actually creates unnecessary care or even harm, termed metric-driven harm. Please describe any experiences you have had with measurement creating unnecessary care or metric-driven harm.
4. Is there any additional information about measurement or measurement-driven behavior you would like to share with us?
5. To which gender identity do you most identify?
6. Which category below includes your age?
7. How long have you been a nurse?
8. Which of the following most accurately describes your current role?
9. In which setting do you work?
10. What is your highest level of education?
11. In which state do you practice as a nurse?
12. How did you hear about our study?
The paradoxical nature of overtreatment stemming from rigid policies, overreliance on standing protocols, and time-bound treatments:
I think time goals are a slippery slope. On the surface, they seem like a great idea, but the reality is, many patients do not fall within textbook cases of diseases, and the rush to meet time goals will lead to unnecessary care and risk to patients.

End-of-life (EOL) treatment in ICU, specifically the withholding of palliative care and continuation of aggressive measures regardless of prognosis, was detailed. This component of overtreatment (invasive and aggressive EOL care) merits special note because it directly fuels undertreatment with comfort measures, with profound impacts on patients and their families. Participants described providers in the ICU neglecting to discuss “code status” as well as prompt interventions a patient or family would want if the patient were to suffer a respiratory or cardiac arrest. This finding is exemplified by the following response from an ICU nurse:
I frequently see patients who deserve palliative care consults that don’t receive them… I believe that those who do not receive palliative care referrals do not receive full informed consent… There are many times that measure-driven [sic] harm has occurred. Unnecessary codes on patients who never received palliative care consults to allow informed consent… Patients suffer regularly until ‘everything has been done’ and patients are now actively dying and are moments from death before they receive comfort measures.

Undertreatment
The paradox of undertreatment and overtreatment was also described by respondents working in non-ICU settings. For example, one respondent noted 20% of patients admitted to her hospital’s observation unit would ultimately be admitted into the hospital inpatient unit. Yet, those patients were placed in a unit with higher nurse-to-patient ratios and fewer resources than those on an inpatient unit. Other illustrations of undertreatment were less subtle (e.g., a participant described a physician withholding intubation in a critically ill patient due to concern over the rate of ventilator-associated pneumonia, a decision that the respondent perceived resulted in the patient’s death). Similarly, a respondent detailed the death of a patient in the ICU that occurred when less-invasive oxygen delivery systems rather than intubation was used, despite the patient’s decreased level of consciousness, worsening pulse oximetry reading, and worsening arterial blood gas. Other illustrations of undertreatment included withholding of a blood transfusion after cardiac surgery due to measurement of the rate of transfusions post-cardiac surgery and premature hospital discharges related to payment-referenced length of stay.

Time-bound treatment metrics were a recurring theme among respondents. The larger repercussions of time-bound treatments were exemplified by an ED nurse who perceived harm related to the “door-to-doctor” metric in her department in relation to hospital throughput and resource utilization. Specifically, in this setting, the largest volume of patients present with symptoms of moderate acuities, such as complaints of abdominal pain, headaches, and flank
pain. Yet, because time-to-treatment is a valued metric, patients with average acuity levels were often seen more quickly than those who were more high-risk, with higher-acuity levels, leaving the latter a far longer time to treatment but an overall departmental average of rapid response.

Undertreatment related to Foley catheter insertions and overreliance on algorithms that exclude the value of clinical observation and judgment was described. To illustrate, a respondent reported an account of a very obese patient with urinary incontinence who was unable to turn in bed without experiencing severe respiratory distress. Her facility’s algorithm determined the patient did not meet the criteria for a catheter, despite the nurse feeling the Foley would benefit the patient’s quality of life. She noted:

Management is more concerned with numbers rather than looking at each patient specifically… Physicians… are not tailoring the plan of care for each specific patient.

**Inefficiency**

The third theme centered around inefficiency. According to the Institute of Medicine (now called the National Academy of Science), “efficient” care involves avoiding the waste of equipment, supplies, ideas, and energy (Agency for Healthcare Research and Quality, 2018). Several participants described nursing care essentially being wasted as nurses followed protocols and policies related in some way to performance measurement. One nurse, for example, described two nurses being required for Foley catheter insertions and two nurses along with either a charge nurse or manager in the room for central line dressing changes, which this nurse felt tied up resources and staff unnecessarily and thus impacted availability of these nurses to other patients. Another participant expressed frustration over nurses having to perform extra, low-value work to improve patient satisfaction scores, despite low staffing ratios. Several of the participants described time-consuming, tedious documentation requirements taking time away from patient care, as exemplified in the following quote.

> On the surface, it seems to clearly be a good idea. But the amount of additional time to document specifics is time consuming and the variability between different insurers or other groups is extremely confusing.

**Discussion**

Although participants were aware of some commonly used metrics, many key metrics were not mentioned. For example, the HVBP requires the measurement and reporting of 20 metrics. Given that performance on these measures may result in a loss of up 2% of revenue, most hospitals focus on high performance in these measures. Yet, a large number of these measures, including the rates of *Clostridium difficile* infection, methicillin-resistant *Staphylococcus aureus* bacteremia, surgical-site infection in colon surgeries and abdominal hysterectomies, total hip and knee arthroplasties, and elective deliveries before 39 weeks of gestation (Centers for Medicare & Medicaid Services, 2017) were not mentioned by any participants in this study. Nor did any participant specifically describe the relationship between metrics and payment. Conversely, participants clearly described inefficiencies and deployment of valuable nurse resources to processes or outcomes that were being measured at the expense of those not measured.

The paradox of overtreatment and undertreatment, creating delivery inefficiencies and nurse frustration, was evident in these findings. Moreover, “tunnel vision” and “measure fixation,” as defined in Smith’s (1995) classic work, was manifest in these findings. Tunnel vision refers to the prioritization of care that is measured over other valuable care or prioritization of measured elements of care over unmeasured care of equal or greater value. Measure fixation is a focus on a particular measurement without reflection on how maximization of the specific outcome can miss or even be in opposition to the underlying objective of care. It also refers to focus on a specific measurement without attention to the distress caused to a patient by maximization of efforts to reach a particular benchmark. Smith posits that ongoing monitoring and evolution of metrics by those closest to the working surface of the system is a crucial strategy toward enhancing long-term, positive outcomes of measurement. In the U.S. healthcare system, nurses, who constitute the most significant percentage of providers and
work across settings (American Association of Colleges of Nursing, 2019), represent precisely that entity. These findings suggest greater awareness of the scope and intent of these measures, and systematic preparation in measurement science, would be a positive step toward that goal.

In this study, nearly all (88%) nurses involved identified harmful or adverse effects witnessed on the frontlines of patient care with the use of quality and performance measures, while only slightly more than half reported positive impacts. There was also some frustration with “management,” but less evidence of participants’ understanding of the financial and reporting requirements driving performance-driven cultures or the intersection of these with nurse leaders’ responsibility to meet these measures. This may be addressed through formalized on-site education.

Nurse leaders have the opportunity to provide context to nurses’ daily work through clarification of the rationale behind performance measures — improved quality at a lower cost — and the specific nuances of different performance-based measurement systems. Moreover, the model emerging from this preliminary study may be used to help nurse and health system leaders maximize the positive outcomes of measurement and minimize — optimally eliminate — measurement-driven harm. Review of the implications of standing orders and algorithms, for example, may serve to reinforce the value of some and lead to modification or elimination of others. The examination of the paradox of overtreatment and undertreatment stemming from measurement also has the potential to support evolution toward a consistently high-performing organization in measures of excellence and, more importantly, in actual reality.

Such system-wide education on the context of measurement and nurse involvement in refinement may lead to enhanced outcomes, less nurse frustration, and evolution of more nuanced measures. Nurse leaders can shape a culture of data fluency, whereby nurses and other healthcare professionals know how to access, mine, and communicate data, and turn it into something meaningful within the healthcare setting (Gemignani et al., 2014). Only by gaining awareness and knowledge regarding the intention and effects of these measures can nurses systematically work with other healthcare providers to ensure the care provided aligns with patients’ best interests while also meeting their organization’s measurement and payment realities.

Moreover, Welton (2016) suggested that, because hospitals’ reimbursement is increasingly referenced to performance measures, measurement of individual nurse performance will soon follow. Others have offered meaningful ways to do so (Yakusheva et al., 2020). Nurse leaders who proactively educate their staff on the parameters and perturbations of measurement will shape an organization that is better prepared for tracking individual nurse performance on nurse-sensitive outcomes as well as nurses’ contributions to overall organizational performance.

The performance and quality metrics currently being used have faced criticism from large bodies of physicians, such as the American College of Physicians, who have argued the measures are mostly invalid and not meaningful in the application of quality, clinical care (Maclean et al., 2018). Growing research from behavioral economists and social psychologists suggests that rewards, such as the financial incentives offered to hospitals and providers who meet performance goals, may dampen motivation and worsen overall performance (Himmelstein et al., 2014). Yet, other studies have found incentive-related improvements (Hoffman & Yakusheva, 2020). The national push for greater transparency in healthcare costs and outcomes (Castellucci, 2019b) ensures measurement will continue to be a part of the contemporary healthcare landscape.

Limitations and Suggestions for Further Research

As a preliminary, hypothesis-generating qualitative study, this study has limited generalizability. Moreover, snowball sampling creates a narrow swath of respondents, a limitation that was intended to be addressed by extensive advertising for informants on Facebook and an online nursing resource website. Nevertheless, these findings suggest that further study regarding the application of performance measures, as witnessed by those most closely involved in the provision of clinical care,
represents an ethical and economic imperative.

**Conclusion**

These findings suggest that more robust nurse understanding of the comprehensive nature of performance measures may allow for continued analysis and improvement of such measures. This, in turn, may lead to improvements in efficiency and patient outcomes and a reduction in measurement-related harm and waste.

---

**References**


Castellucci, M. (2019a, November). Twenty years after To Err is Human hospital care measure are still of little use. Modern Healthcare.


Chambers, M., Bliss, K., & Rambur, B. (2020). Recruiting research participants via traditional snowball vs Facebook advertisements and a website. Western Journal of Nursing Research. https://doi.org/10.1177/0193945920944545


