The Effect of Intensive Care Unit Admission Patterns on Mortality-based Critical Care Performance Measures

Ian J. Barbash^{1,2}, Tri Q. Le^{2,3}, Francis Pike², Amber E. Barnato^{2,3,4}, Derek C. Angus^{2,3}, and Jeremy M. Kahn^{1,2,3}

¹Division of Pulmonary, Allergy, and Critical Care Medicine; ²CRISMA Center, Department of Critical Care Medicine, and ⁴Division of General Internal Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; and ³Department of Health Policy & Management, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania

Abstract

Rationale: Current mortality-based critical care performance measurement focuses on intensive care unit (ICU) admissions as a single group, conflating low-severity and high-severity ICU patients for whom performance may differ and neglecting severely ill patients treated solely on hospital wards.

Objectives: To assess the relationship between hospital performance as measured by risk-standardized mortality for severely ill ICU patients, less severely ill ICU patients, and severely ill patients outside the ICU.

Methods: Using a statewide, all-payer dataset from the Pennsylvania Healthcare Cost Containment Council, we analyzed discharge data for patients with nine clinical conditions with frequent ICU use. Using a validated severity-of-illness measure, we categorized hospitalized patients as either high severity (predicted probability of in-hospital death in top quartile) or low severity (all others). We then created three mutually exclusive groups: high-severity ICU admissions, low-severity ICU admissions, and high-severity ward patients. We used hierarchical logistic regression to generate hospital-specific 30-day risk-standardized mortality rates for each group and then compared hospital performance across groups using Spearman's rank correlation.

Measurements and Main Results: We analyzed 87 hospitals with 22,734 low-severity ICU admissions (mean per hospital, 261 ± 187), 10,991 high-severity ICU admissions (mean per hospital, 126 ± 105), and 6,636 high-severity ward patients (mean per hospital, 76 ± 48). We found little correlation between hospital performance for high-severity ICU patients versus low-severity ICU patients ($\rho = 0.15$; $P = 0.17$). There were 29 hospitals (33%) that moved up or down at least two quartiles of performance across the ICU groups. There was weak correlation between hospital performance for high-severity ICU patients versus high-severity ward patients ($\rho = 0.25$; $P = 0.02$). There were 24 hospitals (28%) that moved up or down at least two quartiles of performance across the high-severity groups.

Conclusions: Hospitals that perform well in caring for high-severity ICU patients do not necessarily also perform well in caring for low-severity ICU patients or high-severity ward patients, indicating that risk-standardized mortality rates for ICU admissions as a whole offer only a narrow window on a hospital's overall performance for critically ill patients.

Keywords: critical care; health services; health care quality assessment; intensive care; patient outcome assessment

(Received in original form September 28, 2015; accepted in final form February 1, 2016)

Supported by National Institutes of Health grants T32 HL007563 (I.J.B.), R01 HL126694 (J.M.K.), and R01 AG035112 (A.E.B.).

Author Contributions: Study concept and design: all authors; data analysis: I.J.B., T.Q.L., and J.M.K.; drafting and revision of the manuscript: I.J.B. and J.M.K.; and critical appraisal and approval of the manuscript version for submission: all authors.

Correspondence and requests for reprints should be addressed to Jeremy M. Kahn, M.D., M.S., Critical Care and Health Policy & Management, University of Pittsburgh, Scaife Hall Room 602-B, 3550 Terrace Street, Pittsburgh, PA 15221. e-mail: kahnjm@upmc.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Ann Am Thorac Soc Vol 13, No 6, pp 877–886, Jun 2016

Copyright © 2016 by the American Thoracic Society

DOI: 10.1513/AnnalsATS.201509-645OC

Internet address: www.atsjournals.org

Measurement of hospital performance for critically ill patients is increasingly important to clinicians, hospital administrators, and health care policy makers. For example, many multicenter

efforts to measure and improve hospital quality rely on critical care performance assessment (1–5), and payment incentives designed to improve quality frequently target the critically ill population (6–10).

These programs typically assess performance using risk-adjusted mortality for patients admitted to an intensive care unit (ICU). Mortality is an important patient-centered outcome, and for decades

the field of critical care has developed and refined methods for detailed clinical risk adjustment (11–14), making risk-adjusted mortality for ICU patients a natural and intuitive quality measure.

Yet, despite the near ubiquity of risk-adjusted ICU mortality as a tool for critical care benchmarking, this measure has a number of conceptual limitations that may limit its utility in hospital performance assessment. First, it includes only patients admitted to an ICU, neglecting the fact that, in the era of ICU outreach teams, much of what is considered to be modern critical care occurs on hospital wards outside the ICU (15–19). Second, it treats ICU patients as a single group, when in reality the factors driving outcomes for low-risk ICU patients, for whom care is focused on observation and prevention of complications (20), may differ from those for high-risk acutely ill patients, for whom care may be focused on urgent resuscitation (21, 22).

These issues are particularly problematic, given variation in ICU admission patterns across hospitals (23–25). Since hospitals differ widely on the basis of both their propensity to treat high-risk patients outside the ICU (26, 27) and their propensity to admit low-risk patients to the ICU (20), a measurement framework based on mortality for all patients admitted to the ICU may provide limited insight into overall hospital performance for the critically ill. To empirically evaluate this issue, we used a statewide database of both ICU and ward patients to assess variation in hospital-level performance assessment when measured on different “types” of critically ill patients: low-severity patients in the ICU, high-severity patients in the ICU, and high-severity patients receiving care solely on the hospital ward. Some of the results of this study have been reported previously in the form of an abstract (28).

Methods

Study Design and Data Sources

We conducted a cross-sectional study using hospital discharge data spanning the period from April 1, 2009, to March 31, 2010, in the Pennsylvania Health Care Cost Containment Council (PHC4) database. The PHC4 database contains detailed patient-level administrative claims records for all hospital discharges in Pennsylvania. Using unique hospital identifiers, we linked

the PHC4 database to the Centers for Medicare and Medicaid Services Healthcare Cost Report Information System to obtain hospital-level data elements.

Unlike most administrative data, the PHC4 dataset contains data on severity of illness upon hospital admission determined using a clinically derived mortality prediction instrument (MediQual Atlas; Quantros, Marlborough, MA), making it a unique resource for this study. MediQual uses laboratory and demographic variables from the day of admission (i.e., blood gases, lactate, troponin, and standard chemistries and complete blood counts). MediQual has prediction accuracy comparable to that of the Acute Physiology and Chronic Health Evaluation II score (29) and significantly improves the accuracy of hospital mortality prediction in comparison to models based solely on administrative data elements (30). Additional information on the MediQual risk predictor is provided in Tables E2 and E3 in the online supplement.

Patients and Variables

We restricted our analysis to patients with 1 of 31 conditions for which PHC4 requires that hospitals report MediQual data. To ensure sufficient numbers of ICU and ward patients, we restricted the analysis to patients admitted with diagnoses for which ICU admission was common, which we defined as at least 33% of patients with that diagnosis. We therefore analyzed patients with one of nine conditions collapsed into five clinical categories: cardiac/vascular surgery (endovascular repair of abdominal aortic aneurysm, open repair of abdominal aortic aneurysm, open or endovascular treatment of carotid vascular disease), cardiac medical (acute myocardial infarction with angioplasty or stent, acute myocardial infarction without angioplasty or stent), respiratory failure (respiratory failure with mechanical ventilation, respiratory failure without mechanical ventilation), sepsis (septicemia), and neurologic (hemorrhagic stroke). These conditions were defined using International Classification of Diseases version 9.0, Clinical Modification, diagnosis and procedure codes (31). We excluded patients without these conditions, patients younger than 16 years of age on the date of admission, and patients missing information on key variables (<1% of sample). To avoid interdependence of observations, if a patient was admitted

more than once in the study period, we included only the initial hospitalization.

We categorized all patients as either high severity or low severity, defining high-severity patients as those in the top quartile of expected in-hospital mortality based on MediQual risk scores and low-severity patients as all others. We then created three mutually exclusive groups: high-severity ICU patients, low-severity ICU patients, and high-severity patients admitted solely to the wards. We identified ICU admissions using standard ICU-specific revenue codes (32). Under this definition, step-down units would be classified as wards, not ICUs. Patients who were admitted to the ICU at any time during their admission were included as ICU patients. We excluded hospitals with fewer than 25 patients in each group, consistent with other studies of hospital performance measurement (33, 34).

The primary outcome variable was 30-day mortality from the date of hospital admission. We obtained postdischarge mortality rates by linking the PHC4 database to the Pennsylvania Department of Health Vital Records file. We chose 30-day mortality because in-hospital mortality may be biased by differing discharge practices across hospitals (35).

Risk adjustment variables included age, sex, comorbidities in the manner of Elixhauser and colleagues (36), admission source (direct, emergency department, or transfer from another acute care hospital), use of mechanical ventilation (32, 37), eligibility diagnosis, primary payer (38), and MediQual predicted risk of in-hospital death. In this capacity, MediQual functions as a severity-of-illness measure rather than as a stand-alone risk predictor, as described previously (35, 39, 40).

Analysis

General analyses. We examined demographic and clinical variables across the three study groups using standard summary statistics. We tested differences between groups using *t* tests for continuous variables and χ^2 tests for categorical variables. We focused on two comparisons: high-severity ICU patients versus low-severity ICU patients and high-severity ICU patients versus high-severity ward patients.

To characterize variation in admission patterns across hospitals, we made two bar graphs: the number of high-severity ICU patients as a proportion of all ICU patients

and the number of high-severity ICU patients as a proportion of all high-severity patients, ordering hospitals by the proportion of high-severity ICU patients. We calculated the median and interquartile ranges for these proportions.

Calculating hospital-specific, risk-adjusted mortality rates. To characterize differences in hospital performance across the three groups, we used hierarchical random effects multivariable regression with patient-level data to generate hospital-level risk-standardized mortality rates (RSMRs) for each of the three groups of patients (41–43). First, we fit a base logistic regression model in which the dependent variable was 30-day mortality and the independent variables were age as quadratic splines, sex, admission source (as indicator covariates), comorbidities (as indicator covariates), eligibility diagnoses as indicator covariates, mechanical ventilation, primary payer, and MediQual predicted risk of death as quadratic splines. We used this model to calculate expected mortality rates for each hospital. Next, we fit a similar model, but with a hospital-specific random effect, using this model to calculate predicted mortality rates for each hospital. Finally, we calculated hospital-specific RSMR as the ratio of predicted mortality to expected mortality multiplied by the statewide mean. We repeated this approach for each patient group, generating three RSMRs for each hospital.

The resulting RSMRs are both risk and reliability adjusted in that they account for instability of the absolute mortality in low-volume hospitals. A small change in the absolute number of patients can translate into dramatic changes in the observed mortality rate that may or may not reflect changes in hospital quality. For hospitals with larger case volumes, the RSMR approximates risk-adjusted mortality. However, for small hospitals with unstable and less reliable observed mortality rates, the RSMR “shrinks” the mortality rate toward the population mean. This modeling approach is identical to those used by the Centers for Medicare and Medicaid Services in their quality reporting programs, as well as to approaches described in the emerging literature on hospital performance assessment (34, 42–44).

Assessing domains of critical care quality. To determine the degree to which each RSMR correlated with the others (and

thus represented a similar domain of quality), we compared hospital RSMRs using Spearman’s rank order correlation coefficient and calculated weighted κ -statistics for interquartile agreement. Again, we focused on two major comparisons: high-severity ICU patients versus low-severity ICU patients and high-severity ICU patients versus high-severity ward patients.

We conducted a sensitivity analysis to test the robustness of our results to the definition of a high-severity patient, changing the definition of high-severity patients to those in the top decile of MediQual-predicted in-hospital mortality. As before, we restricted the analysis to hospitals with at least 25 patients in each group; thus, we performed this analysis on a smaller number of hospitals. Using the same hierarchical mortality prediction models, we generated RSMRs for each group of patients and compared hospitals using Spearman’s rank order correlation.

We performed all analyses using SAS version 9.0 (SAS Institute, Cary, NC) and STATA version 13.1 (StataCorp, College Station, TX) software. We considered a P value less than or equal to 0.05 to be significant. Because deidentified data were used, the study was considered exempt from human subjects review by the University of Pittsburgh Institutional Review Board.

Results

Patient Selection and Characteristics

Of 171 hospitals, 87 had at least 25 patients in each of the three groups and were included in the primary analysis (Figure 1). Within these hospitals, there were 22,734 low-severity ICU admissions (mean per hospital, 261 ± 187), 10,991 high-severity ICU admissions (mean per hospital, 126 ± 105), and 6,636 high-severity ward admissions (mean per hospital, 76 ± 48).

The characteristics of the three patient groups are shown in Table 1. When we compared high-severity ICU patients with low-severity ICU patients, we found that high-severity patients were older, had fewer comorbidities, were more likely to have been admitted from the emergency department, and were more likely to have received mechanical ventilation. In comparison with high-severity ward patients, high-severity ICU patients were

younger, had fewer comorbidities, and were more likely to have received mechanical ventilation. Figure 2 shows the distribution of the proportional relationship between these groups across hospitals, ordered by proportion of high-severity ICU patients. Figure 2 demonstrates wide variation in admission patterns across hospitals with respect to the proportion of ICU patients who were high severity (median, 58%; interquartile range, 50–70%) (Figure 2A) and the proportion of high-severity patients admitted to the ICU versus the ward (median, 32%; interquartile range, 26–37%) (Figure 2B).

Comparison by Hospital Performance

There was poor correlation of hospital performance by RSMR for the three groups of patients (Figure 3). There was no significant correlation between hospital RSMR for low-severity ICU patients versus high-severity ICU patients ($\rho = 0.15$; $P = 0.17$) (Figure 3A). Four hospitals (4.6%) were ranked in the top quartile of performance for high-severity ICU patients but in the bottom quartile of performance for low-severity ICU patients, while three hospitals (3.4%) were ranked in the top quartile of performance for low-severity ICU patients but in the bottom quartile of performance for high-severity ICU patients. A total of 29 hospitals (33%) changed at least two performance quartiles across the two groups of patients, with no statistically significant interquartile agreement ($\kappa = 0.08$; $P = 0.16$).

There was a statistically significant but relatively weak correlation between hospital RSMR for high-severity ICU patients versus high-severity ward patients ($\rho = 0.25$; $P = 0.02$) (Figure 3B). Two hospitals (2.3%) were ranked in the top quartile of performance for high-severity ICU patients but in the bottom quartile of performance for high-severity ward patients, while four hospitals (4.6%) were ranked in the top quartile of performance for high-severity ward patients but in the bottom quartile of performance for high-severity ICU patients. A total of 24 hospitals (28%) changed at least two performance quartiles across the two groups of patients, with very poor interquartile agreement ($\kappa = 0.19$; $P = 0.008$).

Sensitivity Analysis

With a MediQual cutoff of the 90th percentile to define high-severity patients,

Figure 1. Patient selection schema. DRG = diagnosis-related group; ICU = intensive care unit.

there were 28 hospitals with at least 25 patients in each group. There were no statistically significant correlations of hospital RSMR for the three groups of patients: low-severity ICU patients versus high-severity ICU patients ($\rho = -0.06$; $P = 0.75$) or high-severity ICU patients versus high-severity ward patients ($\rho = 0.12$; $P = 0.53$).

Discussion

In a study of Pennsylvania hospitals, we found poor correlation between hospital rankings by risk-adjusted 30-day mortality across three mutually exclusive groups: low-severity ICU patients, high-severity ICU patients, and high-severity ward patients. Hospitals that were top performers for high-severity ICU patients were not necessarily top performers for low-severity ICU patients or high-severity ward patients. These findings suggest the potential for bias

in hospital performance assessments—a phenomenon we term *triage bias*, since varying patterns of ICU triage can lead to varying conclusions about a hospital's performance in the care of critically ill patients.

These findings have important implications for local, regional, and national quality improvement initiatives. Hospitals and health systems across the United States and other countries use risk-adjusted mortality for patients admitted to the ICU as a hospital benchmarking and quality measure (1–5). Our study suggests that risk-adjusted mortality for patients admitted to an ICU may be simultaneously too narrow and too broad to comprehensively measure hospital performance in the care of critically ill patients—too narrow in that it misses important critical care on the hospital ward and too broad in that it fails to distinguish important domains of quality within the ICU.

The lack of correlation between hospital performance for low-severity and high-severity ICU patients suggests that there may be unique domains of critical care quality in the care of these different patient groups, despite the fact that all are in the ICU. For example, hospitals that perform well for low-severity patients may be adept at limiting iatrogenic complications of ICU care such as procedural complications or hospital-acquired infections, which could have a greater impact on the mortality of patients initially at a low risk of death. Conversely, hospitals that perform better for high-severity ICU patients may be adept at acute volume resuscitation and early antibiotic administration for septic shock (45) and may have refined protocols for initiating neuromuscular blockade and prone positioning in patients with severe acute respiratory failure (46, 47).

Similarly, the lack of correlation between hospital performance for high-severity ICU versus high-severity ward

Table 1. Patient characteristics

Characteristic	Low-Severity ICU (n = 22,735)	High-Severity ICU (n = 10,991)	High-Severity Ward (n = 6,636)	Low-Severity ICU vs. High-Severity ICU	High-Severity ICU vs. High-Severity Ward
Age, yr	65.3 ± 14.5	73.9 ± 13.1	81.3 ± 10.6	<i>P</i> < 0.01	<i>P</i> < 0.01
Female sex	9,888 (43.5)	5,469 (49.8)	3,738 (56.3)	<0.01	<0.01
Race				<0.01	<0.01
White	19,041 (83.8)	8,966 (81.6)	5,957 (89.8)		
Black	2,057 (9.0)	1,190 (10.8)	477 (7.2)		
Other	1,637 (7.2)	835 (7.6)	202 (3.0)		
Admission source				<0.01	<0.01
Direct	4,837 (21.3)	1,235 (11.2)	741 (11.2)		
ED	13,991 (61.5)	7,429 (67.6)	4,669 (70.4)		
Other hospital	2,317 (10.2)	1,063 (9.7)	133 (2.0)		
Other	1,590 (7.0)	1,264 (11.5)	1,093 (16.5)		
Comorbidity count				<0.01	<0.01
0	2,963 (13.0)	776 (7.1)	288 (4.3)		
1	6,169 (27.1)	2,693 (24.5)	1,240 (18.7)		
2	6,657 (29.3)	3,553 (32.3)	2,064 (31.1)		
3+	6,946 (30.6)	3,969 (36.1)	3,044 (45.9)		
Comorbidities					
CHF	2,074 (9.1)	2,255 (20.5)	1,513 (22.8)	<0.01	<0.01
COPD	4,229 (18.6)	2,014 (18.3)	1,226 (18.5)	0.54	0.80
Diabetes mellitus	4,193 (18.4)	1,424 (13.0)	1,052 (15.9)	<0.01	<0.01
Liver disease	269 (1.2)	320 (2.9)	130 (2.0)	<0.01	<0.01
Metastatic cancer	214 (0.9)	620 (5.6)	578 (8.7)	<0.01	<0.01
Other cancer	343 (1.5)	413 (3.8)	299 (4.5)	<0.01	<0.01
Primary insurance				<0.01	<0.01
Medicare	13,298 (58.5)	8,572 (78.0)	5,895 (88.8)		
Medicaid	2,492 (11.0)	713 (6.5)	169 (2.5)		
Commercial	2,491 (11.0)	621 (5.7)	251 (3.8)		
Other	4,454 (19.6)	1,085 (9.9)	321 (4.8)		
Mechanical ventilation	4,388 (19.3)	5,870 (53.4)	453 (6.8)	<0.01	<0.01
MediQual predicted risk of death	0.060 ± 0.049	0.397 ± 0.201	0.312 ± 0.144	<0.01	<0.01
Discharge location				<0.01	<0.01
Home	11,905 (52.4)	1,056 (9.6)	867 (13.1)		
SNF	3,105 (13.7)	2,591 (23.6)	2,384 (35.9)		
LTAC	268 (1.2)	317 (2.9)	90 (1.4)		
Dead or hospice	1,941 (8.5)	4,899 (44.6)	2,089 (31.5)		
Other/unknown	5,516 (24.3)	2,128 (19.4)	1,206 (18.2)		
Primary diagnosis					
Cardiovascular surgery	2,997 (13.2)	48 (0.4)	22 (0.3)	<0.01	0.28
Cardiac medical	9,938 (43.7)	975 (8.9)	534 (8.0)	<0.01	0.06
Respiratory	4,143 (18.2)	2,976 (27.1)	1,403 (21.1)	<0.01	<0.01
Septicemia	4,450 (19.6)	5,709 (51.9)	4,086 (61.6)	<0.01	<0.01
Stroke, hemorrhagic	1,207 (5.3)	1,283 (11.7)	591 (8.9)	<0.01	<0.01

Definition of abbreviations: CHF = congestive heart failure; COPD = chronic obstructive pulmonary disease; ED = emergency department; ICU = intensive care unit; LTAC = long-term acute care; SNF = skilled nursing facility.

Values are mean ± SD or frequency (percent). *P* values were calculated with the χ^2 tests for categorical variables and *t* tests for continuous variables.

patients suggests that there may be unique domains of quality for these two patient groups, despite the fact that all are at relatively high risk of mortality. Hospitals that perform well for high-severity patients on the medical ward may be adept at the use of early warning systems, rapid response teams, or cardiac arrest teams (48), care processes that may not necessarily lead to better care in the ICU. Although the literature suggests that rapid response teams and early warning systems have inconsistent effects on mortality (49), it is

likely that many hospitals have developed strategies to use them effectively.

Our findings also highlight the importance of limitations of care in mortality-based hospital benchmarking. The cohort of high-severity patients cared for solely on the hospital ward likely represents a heterogeneous group, some of whom were managed on wards for medical reasons but some of whom were managed on the wards because they elected against aggressive treatment in the ICU. For this latter group of patients, who are more likely

to have multiple comorbidities and active malignancy, high-quality, patient-centered care may still result in death. Thus, a hospital with high-quality critical care services and high-quality palliative care services may have simultaneously lower-than-expected mortality for severely ill patients in the ICU and higher-than-expected mortality for severely ill patients outside the ICU. Indeed, prior literature suggests there is significant variation across hospitals in the tendency to implement early do-not-resuscitate (DNR)

Figure 2. Relative distribution of (A) low-severity versus high-severity intensive care unit (ICU) patients and (B) high-severity ICU versus high-severity ward patients. Vertical axis represents the percentage of patients in each group. Hospitals are ranked from left to right according to increasing percentage of high-severity ICU patients. Each bar represents a single hospital.

orders, independent of measurable patient characteristics (50). These issues reflect a broader limitation of risk-adjusted mortality as an outcome measure in that it inappropriately penalizes hospitals for

providing high-quality end-of-life care (51). To the degree that our results reflect variation in end-of-life practices across hospitals, our study further highlights the need to develop mortality-based

performance measures that are robust to variation in end-of-life care.

Ideally, novel mortality-based benchmarking programs would incorporate patient preferences and care limitations as a

Figure 3. Correlation between (A) hospital risk-standardized mortality rate (RSMR) for high-severity intensive care unit (ICU) patients and hospital RSMR for low-severity ICU patients and (B) hospital RSMR for high-severity ICU patients and hospital RSMR for high-severity ward patients. Each data point represents an individual hospital. Hospitals in the top quartile of performance for one patient group and the bottom quartile of performance for the other patient group (i.e., those moving up or down three quartiles of performance) are shaded in *red*. Hospitals moving up or down two quartiles of performance are shaded in *blue*.

component of risk adjustment models. However, there are several challenges with such an approach. Although most patients who die in the hospital have a DNR order at the time of death (52, 53), the timing of the decision to limit care is difficult to capture in administrative data, and documentation practices may be inconsistent within and across hospitals and different medical record systems. Additionally, there is significant variation across hospitals in the degree to which documentation of treatment preferences and limitations translates into limitations on the care that patients ultimately receive (54), and there is concern that including DNR status in risk adjustment models would simply introduce a new set of biases in hospital benchmarking (55). Determining how best to capture and incorporate patients' treatment preferences and into hospital profiling efforts is a critical topic for future research.

Ultimately, our findings highlight the need for more nuanced quality "fingerprints" that provide actionable information for specific quality improvement opportunities. For example, quality measures that provide information on outcomes for subpopulations of ICU patients would allow ICU directors to allocate resources more efficiently via targeted quality improvement. Quality measures that provide information on high-severity patients outside the ICU would account for programs that allow hospitals to anticipate and manage risk of physiologic deterioration on the ward or to identify and treat patients with prior limitations on therapy in a non-ICU setting.

Our results also suggest a need for composite quality measures that address these different quality domains simultaneously. Empirically derived composite quality measures, which combine different aspects of patient outcomes into a single measure, appear to increase the precision of outcome measurement in comparison with risk-adjusted mortality

alone in general surgical and medical populations (56–58). Similarly structured composite measures could address the limitations of current mortality-based ICU performance measures by simultaneously providing hospitals with a comprehensive assessment of overall performance while allowing for targeted quality improvement interventions for domains in which they need to improve. Ideally, we would derive these measures from the detailed electronic datasets of large health systems, which would allow for more accurate risk adjustment models. For example, one could design a measure that includes variables such as mortality for mechanically ventilated patients, mortality for patients with sepsis or cardiac disease not admitted to the ICU, and hospital-wide ICU admission rates. By modeling the weighting of these variables and their ability to predict future mortality for severely ill patients, such a composite measure could provide feedback to hospital administrators regarding their care of patients at high risk of mortality, those at lower risk of mortality, and overall hospital resource use. From a policy standpoint, these measures would reduce the impact of interhospital variation in triage bias on performance assessment, which may be particularly relevant as payers implement critical care pay-for-performance programs.

Our work has several limitations. First, although we used granular clinical risk adjustment variables, other variables in our study were administratively derived and may have been subject to coding error. Second, because MediQual is a proprietary risk adjustment tool, we do not have access to the specific weights and coefficients for the variables from which it is composed. However, the MediQual risk adjustment system has been used in the published literature examining ICU outcomes (35, 39, 40), and it is unlikely to be systematically less accurate than other risk adjusters in either ICU or hospitalized patients. Third, we examined only a subset of ICU admissions—those with MediQual

assessments and one of the nine clinical conditions we selected. There is a necessary tension between internal validity and generalizability. Although restricting the cohort strengthens our study's internal validity, it may limit our study's external validity, such that our results may not be generalizable to patients with conditions other than those we selected. Despite this consideration, given that the cohorts included patients with myocardial infarction, respiratory failure, and sepsis, our findings are likely to be relevant to many if not all critically ill patients. Fourth, the dataset did not permit an evaluation of performance at the level of individual ICUs. We considered limiting the analysis to one or two conditions typically treated in a single ICU, although most hospitals have a single ICU, and there is often mixing of patients across ICUs within multiunit hospitals. Finally, we cannot directly explain the mechanisms underlying the observed discrepancies in hospital performance; future research is needed to investigate how variation in triage practices might influence hospital performance rankings.

Conclusions

Overall, our study highlights a key limitation of the current paradigm for measuring hospital performance in critical care. Given the variation in hospital performance for subpopulations of critically ill patients within and outside the ICU, current ICU-restricted mortality measures may be simultaneously too broad and too narrow to comprehensively capture all relevant domains of quality in critical care. Novel assessment tools, including both empirically derived composite measures and more detailed quality fingerprints, may be necessary to make critical care performance assessment and quality improvement programs more effective and more actionable. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

- 1 Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med* 2006;34:1297–1310.
- 2 Reineck LA, Le TQ, Seymour CW, Barnato AE, Angus DC, Kahn JM. Effect of public reporting on intensive care unit discharge destination and outcomes. *Ann Am Thorac Soc* 2015;12:57–63.
- 3 Render ML, Deddens J, Freyberg R, Almenoff P, Connors AF Jr, Wagner D, Hofer TP. Veterans Affairs intensive care unit risk adjustment model: validation, updating, recalibration. *Crit Care Med* 2008;36:1031–1042.
- 4 Harrison DA, Brady AR, Rowan K. Case mix, outcome and length of stay for admissions to adult, general critical care units in England, Wales and Northern Ireland: the Intensive Care National Audit & Research Centre Case Mix Programme Database. *Crit Care* 2004;8:R99–R111.

- 5 Boffelli S, Rossi C, Anghileri A, Giardino M, Carnevale L, Messina M, Neri M, Langer M, Bertolini G; Italian Group for the Evaluation of Interventions in Intensive Care Medicine. Continuous quality improvement in intensive care medicine. The GiVITI Margherita Project - Report 2005. *Minerva Anestesiol* 2006;72:419-432.
- 6 Khanduja K, Scales DC, Adhikari NKJ. Pay for performance in the intensive care unit—opportunity or threat? *Crit Care Med* 2009;37:852-858.
- 7 Rhee C, Gohil S, Klompas M. Regulatory mandates for sepsis care—reasons for caution. *N Engl J Med* 2014;370:1673-1676.
- 8 Centers for Medicare & Medicaid Services (CMS), HHS. Medicare program; FY 2014 inpatient prospective payment systems: changes to certain cost reporting procedures related to disproportionate share hospital uncompensated care payments. Interim final rule with comment period. *Fed Regist* 2013;78:61191-61197.
- 9 National Quality Forum. NQF endorses pulmonary and critical care measures. 2012 Jul 31 [accessed 2016 Apr 9]. Available from: http://www.qualityforum.org/News_And_Resources/Press_Releases/2012/NQF_Endorses_Pulmonary_and_Critical_Care_Measures.aspx
- 10 Cooke CR, Iwashyna TJ. Sepsis mandates: improving inpatient care while advancing quality improvement. *JAMA* 2014;312:1397-1398.
- 11 Moreno RP, Bauer P, Metnitz PG. Characterizing performance profiles of ICUs. *Curr Opin Crit Care* 2010;16:477-481.
- 12 Kramer AA, Higgins TL, Zimmerman JE. Comparison of the Mortality Probability Admission Model III, National Quality Forum, and Acute Physiology and Chronic Health Evaluation IV hospital mortality models: implications for national benchmarking*. *Crit Care Med* 2014;42:544-553.
- 13 Kuzniec MW, Vasilevskis EE, Lane R, Dean ML, Trivedi NG, Rennie DJ, Clay T, Kotler PL, Dudley RA. Variation in ICU risk-adjusted mortality: impact of methods of assessment and potential confounders. *Chest* 2008;133:1319-1327.
- 14 Zimmerman JE, Kramer AA. A history of outcome prediction in the ICU. *Curr Opin Crit Care* 2014;20:550-556.
- 15 Berwick DM, Calkins DR, McCannon CJ, Hackbarth AD. The 100,000 Lives Campaign: setting a goal and a deadline for improving health care quality. *JAMA* 2006;295:324-327.
- 16 Louriz M, Abidi K, Akkaoui M, Madani N, Chater K, Belayachi J, Dendane T, Zeggwagh AA, Abouqal R. Determinants and outcomes associated with decisions to deny or to delay intensive care unit admission in Morocco. *Intensive Care Med* 2012;38:830-837.
- 17 Stelfox HT, Hemmelgarn BR, Bagshaw SM, Gao S, Doig CJ, Nijssen-Jordan C, Manns B. Intensive care unit bed availability and outcomes for hospitalized patients with sudden clinical deterioration. *Arch Intern Med* 2012;172:467-474.
- 18 Robert R, Reignier J, Tournoux-Facon C, Boulain T, Lesieur O, Gissot V, Souday V, Hamrouni M, Chapon C, Gouello JP; Association des Réanimateurs du Centre Ouest Group. Refusal of intensive care unit admission due to a full unit: impact on mortality. *Am J Respir Crit Care Med* 2012;185:1081-1087.
- 19 Winters BD, Weaver SJ, Pfoh ER, Yang T, Pham JC, Dy SM. Rapid-response systems as a patient safety strategy: a systematic review. *Ann Intern Med* 2013;158:417-425.
- 20 Rosenthal GE, Sirio CA, Shephardson LB, Harper DL, Rotondi AJ, Cooper GS. Use of intensive care units for patients with low severity of illness. *Arch Intern Med* 1998;158:1144-1151.
- 21 Simchen E, Sprung CL, Galai N, Zitser-Gurevich Y, Bar-Lavi Y, Gurman G, Klein M, Lev A, Levi L, Zveibil F, et al. Survival of critically ill patients hospitalized in and out of intensive care units under paucity of intensive care unit beds. *Crit Care Med* 2004;32:1654-1661.
- 22 Simchen E, Sprung CL, Galai N, Zitser-Gurevich Y, Bar-Lavi Y, Levi L, Zveibil F, Mandel M, Mnataganian G, Goldschmidt N, et al. Survival of critically ill patients hospitalized in and out of intensive care. *Crit Care Med* 2007;35:449-457.
- 23 Seymour CW, Iwashyna TJ, Ehlenbach WJ, Wunsch H, Cooke CR. Hospital-level variation in the use of intensive care. *Health Serv Res* 2012;47:2060-2080.
- 24 Wunsch H, Wagner J, Herlim M, Chong DH, Kramer AA, Halpern SD. ICU occupancy and mechanical ventilator use in the United States. *Crit Care Med* 2013;41:2712-2719.
- 25 Chen LM, Render M, Sales A, Kennedy EH, Witala W, Hofer TP. Intensive care unit admitting patterns in the Veterans Affairs health care system. *Arch Intern Med* 2012;172:1220-1226.
- 26 Jones D, George C, Hart GK, Bellomo R, Martin J. Introduction of medical emergency teams in Australia and New Zealand: a multi-centre study. *Crit Care* 2008;12:R46.
- 27 Tirkkonen J, Nurmi J, Olkkola KT, Tenhunen J, Hopppu S. Cardiac arrest teams and medical emergency teams in Finland: a nationwide cross-sectional postal survey. *Acta Anaesthesiol Scand* 2014;58:420-427.
- 28 Barbash IJ, Le TQ, Pike F, Barnato AE, Angus DC, Kahn JM. Judging hospitals for care of all critically ill patients versus care of all intensive care unit patients: is there a difference [abstract]? *Am J Respir Crit Care Med* 2015;191:A5201.
- 29 Iezzoni LI. The risks of risk adjustment. *JAMA* 1997;278:1600-1607.
- 30 Pine M, Jordan HS, Elixhauser A, Fry DE, Hoaglin DC, Jones B, Meimban R, Warner D, Gonzales J. Enhancement of claims data to improve risk adjustment of hospital mortality. *JAMA* 2007;297:71-76.
- 31 Pennsylvania Health Care Cost Containment Council. Technical notes for the hospital performance report. December 2014 [accessed 2016 Apr 9]. Available from: <http://www.phc4.org/reports/hpr/13/docs/hpr2013technotes.pdf>
- 32 Quan H, Parsons GA, Ghali WA. Validity of procedure codes in International Classification of Diseases, 9th revision, clinical modification administrative data. *Med Care* 2004;42:801-809.
- 33 Barreto-Filho JA, Wang Y, Rathore SS, Spatz ES, Ross JS, Curtis JP, Nallamothu BK, Normand S-LT, Krumholz HM. Transfer rates from nonprocedure hospitals after initial admission and outcomes among elderly patients with acute myocardial infarction. *JAMA Intern Med* 2014;174:213-222.
- 34 Sjoding MW, Prescott HC, Wunsch H, Iwashyna TJ, Cooke CR. Hospitals with the highest intensive care utilization provide lower quality pneumonia care to the elderly. *Crit Care Med* 2015;43:1178-1186.
- 35 Reineck LA, Pike F, Le TQ, Cicero BD, Iwashyna TJ, Kahn JM. Hospital factors associated with discharge bias in ICU performance measurement. *Crit Care Med* 2014;42:1055-1064.
- 36 Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8-27.
- 37 Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE, Ghali WA. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;43:1130-1139.
- 38 Lyon SM, Benson NM, Cooke CR, Iwashyna TJ, Ratcliffe SJ, Kahn JM. The effect of insurance status on mortality and procedural use in critically ill patients. *Am J Respir Crit Care Med* 2011;184:809-815.
- 39 Kim MM, Barnato AE, Angus DC, Fleisher LA, Kahn JM. The effect of multidisciplinary care teams on intensive care unit mortality. *Arch Intern Med* 2010;170:369-376.
- 40 Barnato AE, Chang CCH, Farrell MH, Lave JR, Roberts MS, Angus DC. Is survival better at hospitals with higher "end-of-life" treatment intensity? *Med Care* 2010;48:125-132.
- 41 Normand SLT, Glickman ME, Gatsonis CA. Statistical methods for profiling providers of medical care: issues and applications. *J Am Stat Assoc* 1997;92:803-814.
- 42 Krumholz HM, Wang Y, Mattera JA, Wang Y, Han LF, Ingber MJ, Roman S, Normand SLT. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with an acute myocardial infarction. *Circulation* 2006;113:1683-1692.
- 43 Krumholz HM, Wang Y, Mattera JA, Wang Y, Han LF, Ingber MJ, Roman S, Normand SLT. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with heart failure. *Circulation* 2006;113:1693-1701.
- 44 Drye EE, Normand SLT, Wang Y, Ross JS, Schreiner GC, Han L, Rapp M, Krumholz HM. Comparison of hospital risk-standardized mortality rates calculated by using in-hospital and 30-day models: an observational study with implications for hospital profiling. *Ann Intern Med* 2012;156:19-26.
- 45 Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, Sevransky JE, Sprung CL, Douglas IS, Jaeschke R, et al.; Surviving Sepsis Campaign Guidelines Committee including the Pediatric

- Subgroup. Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock: 2012. *Crit Care Med* 2013;41:580–637.
- 46 Papazian L, Forel JM, Gacouin A, Penot-Ragon C, Perrin G, Loundou A, Jaber S, Arnal JM, Perez D, Seghboyan JM, *et al.*; ACURASYS Study Investigators. Neuromuscular blockers in early acute respiratory distress syndrome. *N Engl J Med* 2010;363:1107–1116.
 - 47 Guérin C, Reignier J, Richard JC, Beuret P, Gacouin A, Boulain T, Mercier E, Badet M, Mercat A, Baudin O, *et al.*; PROSEVA Study Group. Prone positioning in severe acute respiratory distress syndrome. *N Engl J Med* 2013;368:2159–2168.
 - 48 Barbash IJ, Kahn JM. Organizational approaches to improving resuscitation effectiveness. *Crit Care Clin* 2015;31:165–176.
 - 49 Chan PS, Jain R, Nallmothu BK, Berg RA, Sasson C. Rapid response teams: a systematic review and meta-analysis. *Arch Intern Med* 2010;170:18–26.
 - 50 Zingmond DS, Wenger NS. Regional and institutional variation in the initiation of early do-not-resuscitate orders. *Arch Intern Med* 2005;165:1705–1712.
 - 51 Holloway RG, Quill TE. Mortality as a measure of quality: implications for palliative and end-of-life care. *JAMA* 2007;298:802–804.
 - 52 Jones DA, McIntyre T, Baldwin I, Mercer I, Kattula A, Bellomo R. The medical emergency team and end-of-life care: a pilot study. *Crit Care Resusc* 2007;9:151–156.
 - 53 Holley A, Kravet SJ, Cordts G. Documentation of code status and discussion of goals of care in gravely ill hospitalized patients. *J Crit Care* 2009;24:288–292.
 - 54 Hart JL, Harhay MO, Gabler NB, Ratcliffe SJ, Quill CM, Halpern SD. Variability among US intensive care units in managing the care of patients admitted with preexisting limits on life-sustaining therapies. *JAMA Intern Med* 2015;175:1019–1026.
 - 55 Tabak YP, Johannes RS, Silber JH, Kurtz SG. Should Do-Not-Resuscitate status be included as a mortality risk adjustor? The impact of DNR variations on performance reporting. *Med Care* 2005;43:658–666.
 - 56 Dimick JB, Birkmeyer NJ, Finks JF, Share DA, English WJ, Carlin AM, Birkmeyer JD. Composite measures for profiling hospitals on bariatric surgery performance. *JAMA Surg* 2014;149:10–16.
 - 57 Staiger DO, Dimick JB, Baser O, Fan Z, Birkmeyer JD. Empirically derived composite measures of surgical performance. *Med Care* 2009;47:226–233.
 - 58 Chen LM, Staiger DO, Birkmeyer JD, Ryan AM, Zhang W, Dimick JB. Composite quality measures for common inpatient medical conditions. *Med Care* 2013;51:832–837.