DOI: 10.1111/j.1475-6773.2009.00959.x
RESEARCH ARTICLE

The Relationship between Hospital Volume and Mortality in Mechanical Ventilation: An Instrumental Variable Analysis

Jeremy M. Kahn, Thomas R. Ten Have, and Theodore J. Iwashyna

Objective. To examine the relationship between hospital volume and mortality for nonsurgical patients receiving mechanical ventilation.

Data Sources. Pennsylvania state discharge records from July 1, 2004, to June 30, 2006, linked to the Pennsylvania Department of Health death records and the 2000 United States Census.

Study Design. We categorized all general acute care hospitals in Pennsylvania ($n = 169$) by the annual number of nonsurgical, mechanically ventilated discharges according to previous criteria. To estimate the relationship between annual volume and 30-day mortality, we fit linear probability models using administrative risk adjustment, clinical risk adjustment, and an instrumental variable approach.

Principle Findings. Using a clinical measure of risk adjustment, we observed a significant reduction in the probability of 30-day mortality at higher volume hospitals ($\geq$ 300 admissions per year) compared with lower volume hospitals ($<$300 patients per year; absolute risk reduction: 3.4%, $p = .04$). No significant volume–outcome relationship was observed using only administrative risk adjustment. Using the distance from the patient's home to the nearest higher volume hospital as an instrument, the volume–outcome relationship was greater than observed using clinical risk adjustment (absolute risk reduction: 7.0%, $p = .01$).

Conclusions. Care in higher volume hospitals is independently associated with a reduction in mortality for patients receiving mechanical ventilation. Adequate risk adjustment is essential in order to obtained unbiased estimates of the volume–outcome relationship.

Key Words. Critical care, intensive care, respiratory failure, risk adjustment, mortality

Increased case load is associated with improved outcomes in many areas of health care, including trauma, acute myocardial infarction, and many types of high-risk surgeries (Halm, Lee, and Chassin 2002). Recent studies have documented a relationship between volume and outcome in critical care as well

(Kahn 2007). Among patients admitted to an intensive care unit (ICU), increased hospital admission volume is associated with improved survival in those with acute respiratory failure, sepsis, and subsets of patients at a high risk for death (Durairaj et al. 2005; Glance et al. 2006; Kahn et al. 2006; Peelen et al. 2007).

Although the majority of ICU volume–outcome studies have shown a significant relationship, this finding is not universal. At least one population-based study in patients undergoing mechanical ventilation showed no significant volume–outcome relationship (Needham et al. 2006). A possible explanation for this discrepancy is that the positive studies uniformly used detailed clinical risk adjustment, while the negative study used administrative variables for risk adjustment. Clinical risk adjustment may be necessary to fully account for differences in severity of illness, which is an important confounder between volume and outcome (Tsai et al. 2006). Indeed, several studies have shown that the addition of clinical data to administrative data can improve the predictive accuracy of multivariate models (Hannan et al. 1992; Pine et al. 2007). Another possibility is that the negative study was performed on a population sample of hospitals, while most positive studies were not. The existence of a volume–outcome effect in selected hospitals may not reflect the true relationship in broader settings.

The purpose of this study was to examine the relationship between hospital volume and outcome in a large population-based sample of nonsurgical mechanically ventilated patients. We also examined the importance of risk adjustment in ICU volume–outcome studies to determine if differences in risk adjustment could explain the differences in results. We used an administrative data set from the Pennsylvania containing a clinical severity of illness measure (MediQual Atlas), allowing us to compare the volume–outcome effect using clinical and administrative risk adjustment. We also performed an instrumental variable analysis to ascertain whether an unbiased estimate of the volume–outcome effect in mechanical ventilation could be obtained without clinical risk adjustment. We hypothesized that an instrumental variable approach could determine whether the association between volume and outcome is independent of measured and unmeasured confounders and potentially causal in nature.

Address correspondence to Jeremy M. Kahn, M.D., M.Sc., Assistant Professor of Medicine, Division of Pulmonary, Allergy & Critical Care, Center for Clinical Epidemiology & Biostatistics, University of Pennsylvania School of Medicine, 723 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104; e-mail: jmkahn@mail.med.upenn.edu. Thomas R. Ten Have, Ph.D., M.P.H., is with Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, Philadelphia, PA. Theodore J. Iwashyna, M.D., Ph.D., is with Division of Pulmonary & Critical Care, University of Michigan School of Medicine, Ann Arbor, MI.

METHODS

Study Design and Patients

We performed a retrospective cohort study using Pennsylvania state discharge data obtained from the Pennsylvania Health Care Cost Containment Council (PHC4). PHC4 collects claims data on all patients discharged from nonfederal hospitals within the Commonwealth of Pennsylvania. All discharges between July 1, 2004, and June 30, 2006, involving intensive care were eligible for the analysis. We excluded children's hospitals, rehabilitation hospitals, long-term acute care hospitals, and specialty surgery hospitals. Patient-level data were linked to the Pennsylvania Department of Health's death records to obtain each patient's vital status at 30 days after hospital admission. Population and geographic information were obtained by linking the data to the 2000 United States Census by the patient's ZIP code of residence. Hospital characteristics were obtained from the 2005 American Hospital Association Annual Survey.

Patients undergoing mechanical ventilation were identified by the *International Classification of Diseases version 9.0—Clinical Modification* (ICD-9-CM) procedure codes 96.70, 96.71, and 96.72 (mechanical ventilation time unspecified, <96 consecutive hours, and ≥ 96 consecutive hours, respectively). We excluded patients <16 years of age because the causes and outcomes of respiratory failure are fundamentally different between the adult and pediatric populations. We excluded patients undergoing major surgery identified using ICD-9-CM procedure codes in order to eliminate the portion of the volume–outcome relationship potentially attributable to surgery-specific volume. For patients with more than one hospitalization during the study period, we examined only the first hospitalization. Because we planned an instrumental variable analysis using distance as an instrument, it was necessary to examine only patients with a defined home location and a probability of admission to either a low or higher volume hospital (Harris and Remler 1998). Consequently we excluded patients with missing ZIP codes or ZIP codes outside the state, patients admitted directly from other acute care hospitals (interhospital transfers), and patients who traveled more than 75 miles to reach their admission hospital. Patients transferred from study hospitals to other acute care hospitals were retained in the analysis.

Variables

The primary exposure variable was the volume of mechanically ventilated patients for each hospital, annualized over the study period. For hospitals that opened or closed during the study period, we calculated annual volume based

on the length of time the hospital was open. Annual volume was calculated before patient exclusions. We categorized volume into five groups (<100 admissions per year, 100–199 admissions per year, 200–299 admissions per year, 300–599 admissions per year, and ≥ 600 admissions per year) using previously defined cut-points (Needham et al. 2006). This allowed us to directly compare our patient population with previously published studies. To simplify the analysis, we further categorized the exposure into two groups: lower volume hospitals (<300 admissions per year) and higher volume hospitals (≥ 300 admissions per year). The primary outcome variable was mortality at 30 days from hospital admission.

Variables for risk adjustment were determined a priori as potential confounders of the relationship between volume and outcome and included age, gender, admission source, comorbidities, severity of illness on hospital admission, hospital teaching status, and hospital technological capacity. Age, gender, and admission source were directly available in the discharge data set. Comorbidities were determined using ICD-9-CM diagnosis codes from each patient's index admission. We measured comorbidities in two ways: the Deyo modification of the Charlson comorbidity index (Deyo, Cherkin, and Ciol 1992) and select individual Elixhauser comorbidities (cirrhosis, leukemia, lymphoma, and solid cancer) known to be important predictors of mortality in critically ill patients (Elixhauser et al. 1998; Zimmerman et al. 2006). Severity of illness was measured using the MediQual Atlas probability of in-hospital death, a validated risk adjustment tool for hospitalized patients using key clinical and demographic variables measured on the onset of hospitalization (Iezzoni 1997). A MediQual Atlas score is automatically calculated by PHC4 on most patients admitted to Pennsylvania hospitals but may be absent due to missing clinical data. Teaching status was determined from each hospital's resident-to-bed ratio and categorized into nonteaching (ratio = 0), small teaching (ratio between 0 and 0.2), and large teaching (ratio ≥ 0.2).

To determine each hospital's technological capacity, we measured the hospital-specific incidence of nine key procedures during the study period: neurological surgery, cerebral arteriogram, coronary-artery bypass grafting, percutaneous transluminal coronary angioplasty, diagnostic coronary angiography, hemodialysis, magnetic resonance imaging, electroencephalogram, and noncoronary/noncerebral arteriogram or venogram using ICD-9-CM procedure codes. To avoid misclassification due to coding error, we considered hospitals to possess the technology if they coded more than 10 procedures during the 2-year study period. We then created a technological capacity index equal to the weighted sum of the individual technologies, with each

technology weighted by the percentage of hospitals that do not have it, thus giving greater weights to hospitals with technologies that are rare (Landon et al. 2006).

We chose the distance from each patient's home to the nearest higher volume hospital ($\geq$ 300 mechanically ventilated admissions per year) as a potential instrument. We hypothesized that distance to a higher volume hospital will influence a patient's probability of receiving care in a higher volume hospital, and that this is the only mechanism by which distance to a higher volume hospital will influence mortality (Kennedy 2002). Similar distances have been used effectively in several other instrumental variable analyses (McClellan, McNeil, and Newhouse 1994; Stukel et al. 2007). We also examined the differential distance between the patient's home to the nearest small-volume hospital and the patient's home to nearest large-volume hospital, reflecting a model of hospital choice based on differential rather than absolute difference (Beck et al. 2003). Distances were calculated as the linear arc distance between the exact longitude and latitude of the hospital to the latitude and longitude of the ZIP code centroid of the patient's residence.

Analysis

The analyses were based on regression models adjusting for observed confounders and instrumental variable regression adjusting for both observed and unobserved confounding. We used the TREATREG command in *Stata* specifying robust Huber–White standard errors and confidence intervals to account for hospital-level clustering and potential heteroskedasticity.

We used patient-level multivariate linear probability regression to determine the effect of hospital admission volume on 30-day mortality controlling for important confounders. Linear probability models use the probability of mortality as the dependent variable and are easily adaptable to instrumental variable analyses. Although these models do not constrain the probability of death to a number between 0 and 1, this limitation is typically not an issue in mechanically ventilated patients, who have a risk of death around 30 percent (Agresti 2002; Esteban et al. 2002). Separate models were created containing just the primary exposure (base model); the primary exposure plus age, gender, admission source, and comorbidities (administrative model); and the variables in the administrative model plus the MediQual Atlas probability of in-hospital death (clinical model). For the instrumental variable analysis, we performed a two-stage maximum likelihood estimation process, specifying linear models for both the probability of in-hospital death (using the admin-

istrative model) and the probability admission to a high-volume hospital conditional on the instrument and other observed covariates (Maddala 1983).

To assess the assumptions underlying the instrumental variable analysis, we performed the following analyses. First, we used Pearson's correlation coefficient to examine the correlation between the instrument (distance to the nearest large-volume hospital), the primary exposure (admission to a large-volume hospital), and the primary outcome (30-day mortality). Second, we performed a partial *F*-test under a linear regression in which volume group was regressed on distance and a complete set of patient characteristics. With a valid instrument the test statistic should be very high (i.e., distance strongly predicts volume group conditional on patient characteristics). Third, we performed a partial *F*-test under a linear probability model in which mortality was regressed on distance and the complete set of patient characteristics. With a valid instrument the test statistic should be very low (i.e., distance does not predict mortality conditional on patient characteristics and adjusting for volume), recognizing that a lack of association may be confounded by unmeasured variables. Fourth, we compared observable health characteristics between patients closer than the median distance to large-volume hospital and farther than the median distance to a large-volume hospital. For the instrument to be valid, these two patient groups should appear qualitatively similar with respect to the exposure.

To confirm the presence of endogeneity in our primary exposure, we performed the Rivers–Vuong test (Rivers and Vuong 1988) and the Hausman specification test (Hausman and McFadden 1984). The Rivers–Vuong test assesses the relationship between mortality and the residual from the regression of the volume group on distance by including this residual with volume as the covariates in the mortality model. The Hausman specification tests assess whether the least-squares linear probability model and the instrumental variable model are consistent, obviating the need for an instrumental variable analysis.

All analyses were performed using *Stata* 9.2 (College Station, TX). A p value of $\leq .05$ was considered significant. This work was considered exempt from Human Subjects Review by the University of Pennsylvania Institutional Review Board.

RESULTS

There were 169 general acute-care hospitals in the state during the study period, with a total of 78,427 nonsurgical admissions requiring mechanical ventilation.

Table 1: Hospital Characteristics by Annual Volume of Nonsurgical Mechanically Ventilated Patients

	Annual Admission Volume				
Variable	*< 100*	*100–199*	*200–299*	*300–599*	*600+*
Hospitals (*n*)	84	37	25	17	6
Total beds	109 ± 67	221 ± 78	273 ± 80	554 ± 137	818 ± 373
For profit	7 (9)	5 (14)	0 (0)	1 (6.3)	0 (0)
Teaching status					
Nonteaching	56 (75)	18 (50)	6 (26)	0 (0)	0 (0)
Small teaching	16 (21)	15 (42)	13 (57)	7 (44)	2 (33)
Large teaching	2 (4)	3 (8)	4 (17)	9 (56)	4 (67)
Technological capability index*	0.8 ± 1.1	2.6 ± 1.2	2.4 ± 1.3	3.9 ± 0.6	4.5 ± 0.7
MSA size					
<100,000	32 (38)	1 (3)	2 (8)	1 (6)	0 (0)
100,000–1 million	32 (38)	10 (27)	5 (20)	9 (53)	1 (17)
1 million+	20 (24)	26 (70)	18 (72)	7 (42)	5 (83)

Values are either mean ± standard deviation or frequency (%).

*Scale indicating the number of high-technology procedures available at the hospital weighted by the scarcity of the procedure (range: 0.00–5.05). See "Methods" for further explanation.

MSA, metropolitan statistical area.

The median number of mechanical ventilation admissions was 100, with a range from 1 to 1,994. Compared with lower volume hospitals, higher volume hospitals had more hospital beds, were more likely to be teaching hospitals, were located in larger communities, and had higher technological capacity (Table 1).

We excluded 452 patients (0.6%) <16 years of age, 19,780 patients (25.2%) who underwent major surgery, 21,261 repeat admissions (27.1%), 1,964 patients admitted in transfer from another hospital (2.5%), 1,711 patients (2.2%) with missing or bad ZIP codes, and 272 patients (0.4%) who traveled more than 75 miles to reach their admission hospital. The final sample contained 30,677 patients (Table 2). Compared with patients in lower volume hospitals, patients in higher volume hospitals were younger and more likely to be non-white. Comorbidities were similar among all volume groups. The MediQual Atlas probability of in-hospital death was available for 24,726 patients (81%) and was slightly greater at higher volume hospitals. Unadjusted mortality was lower at higher volume hospitals.

Median distance from the patients' residence to the nearest higher volume hospital (>300 patients per year) was 7.6 miles (interquartile range: 2.9–21.2 miles). The nearest higher volume hospital was not the admitting

Table 2: Characteristics of Nonsurgical Patients Undergoing Mechanical Ventilation by Annual Hospital Admission Volume

	Annual Admission Volume				
Variable	*< 100*	*100–199*	*200–299*	*300–599*	*600+*
Patients (*n*)	4,178	5,853	6,431	8,202	6,003
Age	68 ± 17	67 ± 17	65 ± 18	60 ± 18	60 ± 20
Female (%)	49	51	51	45	44
Race (%)					
White	93	88	77	78	61
Black	4	7	19	15	26
Other	3	5	4	7	13
Admission source (%)					
ED	83	83	92	83	85
Direct	17	17	8	17	15
Charlson index					
Mean	1.8 ± 1.6	1.8 ± 1.7	1.8 ± 1.7	1.5 ± 1.6	1.6 ± 1.8
> 3 (%)	10	11	11	9	11
Primary diagnosis (%)					
Respiratory	50	43	43	29	29
Cardiac	14	14	13	10	10
Neurological	7	8	9	21	22
Other	28	35	35	39	40
MedQual probability of in-hospital death*	0.24	0.23	0.24	0.27	0.25
Distance from home to hospital (m)	4.3 [2.4–9.1]	3.6 [2.0–6.8]	3.2 [1.5–6.3]	5.2 [2.6–12.3]	4.6 [1.9–10.5]
Discharge location (%)					
Home	24	23	25	34	34
Other hospital	18	13	11	6	5
SNF/LTAC	21	25	24	23	21
Dead	33	34	34	32	33
Other	5	5	6	6	7
30-day mortality	40.4	39.1	37.0	32.4	32.4

*Available for 24,726 patients (81%).

ED, emergency department; SNF, skilled nursing facility; LTAC, long-term acute care.

hospital for 16,462 patients (53.7%). For 5,757 patients (18.2%) the closest hospital to home was a higher volume hospital. Of patients admitted to a small-volume hospital, 1,840 (11.2%) lived closer to a higher volume hospital than the admitting hospital.

Distance to the nearest higher volume hospital was strongly associated with admission to a higher volume hospital (Pearson's coefficient: -0.39; F statistic from linear probability model = 2000.9, $p < .001$). Conversely,

Table 3: Patient Characteristics by Distance to Nearest Higher Volume Hospital (≥ 300 Patients per Year)

Variable	*< 7.6 Miles (n = 15,357)*	*≥ 7.6 Miles (n = 15,310)*
Age	63 ± 19	64 ± 19
Female (%)	48	47
Race (%)		
White	65	92
Black	27	3
Other	8	5
Median income of census tract, in thousands	35.6 [27.0–44.9]	38.0 [32.4–49.0]
Charlson index	1.6 ± 1.7	1.6 ± 1.7
Major comorbidities (%)		
Cirrhosis	1	1
Leukemia	1	1
Lymphoma	1	1
Other cancer	9	9
Primary admission diagnosis (%)		
Respiratory	35	40
Cardiac	12	12
Neurological	15	14
Other	38	34

Values are either mean ± standard deviation, median [interquartile range], or percent.

distance to the nearest higher volume hospital was uncorrelated with 30-day mortality (Pearson's coefficient: 0.02; F statistic from linear probability model = 1.80, $p = .18$). Health characteristics were similar between patients closer to a higher volume hospital than the median and those farther away from a higher volume hospital than the median (Table 3). The exception was race—black patients were much more likely than white patients to live close to a higher volume hospital. The median income of census tract of residence, however, was similar between distance groups. Similar to our primary instrument, differential distance was also highly correlated with admission to a high-volume hospital (F statistic = 2269.6, $p < .001$) and uncorrelated with 30-day mortality (F statistic = 0.52, $p = .47$).

The relationship between hospital volume and 30-day mortality is shown in Table 4. Comparing patients admitted to a lower volume hospital (<300 patients per year) to patients admitted to a higher volume hospital (≥ 300 patients per year) using only administrative risk adjustment there was no significant mortality benefit at higher volume hospitals (absolute risk

Table 4: Relationship between Annual Hospital Volume and 30-Day Mortality Using a Linear Probability Regression Model*

Model	Δp	*95% CI*	*p-value*
Base model (only volume)	−.062	−0.085, −0.040	<.01
Age and gender	−.015	−0.034, 0.004	.11
Age, gender, income, admission source, comorbidities, teaching status, and technological capability	−.022	−0.053, 0.008	.14
Age, gender, income, admission source, comorbidities, teaching status, technological capability, and MedQual†	−.034	−0.066, −0.012	.04
Age, gender, income, admission source, comorbidities, teaching status, technological capability, and IV (absolute difference)	−.070	−0.125, −0.015	.01
Age, gender, income, admission source, comorbidities, teaching status, technological capability, and IV (differential difference)	−.055	−0.103, −0.007	.03

*Adjusted probability differences are for hospitals with >300 admission per year compared with hospitals with <300 admission per year.

†Available for 24,726 patients (81%).

IV, instrumental variable.

reduction: 2.2%, $p = .14$). In the subset of patients for whom the MediQual Atlas score was available, clinical risk adjustment resulted in a small but statistically significant reduction in the probability of death associated with care at a higher volume hospital (absolute risk reduction: 3.4%, $p = .04$). In the instrumental variable analysis, admission to a higher volume hospital was associated with a large and significant reduction in the risk of death (absolute risk reduction: 7.0%, $p = .01$). Using differential distance as the instrument yielded attenuated but similar results (absolute risk reduction: 5.4%, $p = .03$). The Rivers–Vuong test for endogeneity rejected the null hypothesis ($p = .04$), indicating the presence of indication bias in the administrative model. The Hausman test was highly significant ($p < .001$) indicating that the ordinary least-squares model results in biased estimates compared with the instrumental variable analysis.

DISCUSSION

In medical patients receiving mechanical ventilation, care in a higher volume hospital was independently associated with a significant reduction in the adjusted risk of death compared with care in a small-volume hospital. No

relationship could be detected using only administrative risk adjustment. Using clinical risk adjustment, a statistically significant result was observed, although the point estimate was qualitatively only slightly greater than in the administrative model. Using an instrumental variable approach to control for unmeasured confounding resulted in a clinically and statistically significant relationship between volume and mortality.

Our results help explain discordant findings on this topic in the medical literature. Of the two recent studies examining the volume–outcome relationship in mechanically ventilated patients, one showed a significant relationship and one did not (Kahn et al. 2006; Needham et al. 2006). The positive study used detailed clinical risk adjustment comprised of physiologic variables on day 1 of the patients' ICU admission (Zimmerman et al. 2006). The negative study used administrative risk adjustment available in standard hospital claims. We show that the volume–outcome relationship is extremely sensitive to clinical risk adjustment—the positive result we observed using clinical risk adjustment was not observed using only administrative risk adjustment. An instrumental variable approach produced similar results but resulted in an even greater absolute risk reduction compared with clinical risk adjustment. This finding not only suggests that administrative risk adjustment alone is subject to unmeasured confounding but also suggests that the MediQual Atlas clinical risk adjustment may not account for all the severity differences between large- and small-volume hospitals. Alternatively, the difference in treatment effects may be due to the different types of models. The instrumental variable analysis estimates the marginal treatment effect (the effect on patients whose probability of admission to a high-volume hospital is only due to distance) rather than the average treatment effect (the effect on the entire population of mechanical ventilated patients) (Harris and Remler 1998). Marginal patients may have more potential benefit from admission to a higher volume hospital than average patients.

Our study also demonstrates the utility of instrumental variable analyses for obtaining unbiased estimates of treatment effects in critical care health services research. The results from our instrumental variable analysis are remarkably similar to results from the previous study of mechanically ventilated patients which used high-quality clinical risk adjustment (Kahn et al. 2006). Distance-based instruments are commonly used in health services research to help overcome selection bias and unmeasured confounding (McClellan, McNeil, and Newhouse 1994; Brooks et al. 2003; Tsai et al. 2006; Stukel et al. 2007). In our study, distance to the nearest higher volume hospital was strongly associated with the exposure and uncorrelated with both the outcome

of interest and other observable health characteristics, fulfilling the criteria of a valid instrument. It is possible that distance from home to hospital can be used as an instrument to answer other ICU-related health services research questions that are subject to indication bias. Of some concern was the finding that the instrument was strongly associated with patient race. In some health care settings black race is associated with poorer health status and higher mortality (Harris 2001; Konety, Vaughan Sarrazin, and Rosenthal 2005). However, more recent data suggest that these differences are more often due to socioeconomic factors and differences in treatment patterns (Petersen et al. 2002; Barnato et al. 2005). In patients with critical illness, race has not been shown to be associated with mortality independent of comorbidities and severity of illness (Williams et al. 1995; Dombrovskiy et al. 2007; Barnato et al. 2008).

Instrumental variable analyses are frequently used to determine causal relationships from observational data. Although our primary goal was to determine the association between volume and outcome independent of potential unmeasured confounding, our instrumental variable result suggests a causal relationship between volume and outcome in mechanically ventilated patients. Volume in this case is likely a proxy for clinical experience, which may be causally related to outcome under a "practice makes perfect" conceptual model (Luft, Hunt, and Maerki 1987). As clinicians gain experience in the care of complex patients they may provide better care. Our findings provide new support a causal effect of volume in mechanically ventilated patients. However, we still cannot rule out that other unmeasured hospital level factors correlated with volume are the true causal mediators.

The volume–outcome relationship in critical care has important health policy implications. Variation in quality across ICUs has prompted calls for regionalization of critical care in a manner similar to trauma or neonatal care (Angus and Black 2004; Barnato et al. 2007). Regionalization would involve routinely transferring high-risk ICU patients to a small number of large, high-quality hospitals. If patients transferred to high-volume centers experience similar outcomes to patients originally admitted to those centers, then regionalization has the potential to significantly improve outcomes (Kahn et al. 2008). Regionalization also has the potential to improve efficiency in the ICU by taking advantage of economies of scale (Jacobs, Rapoport, and Edbrooke 2004). The existence of a strong volume–outcome relationship supports continued investigations into the feasibility of developing regionalized systems of critical care.

The volume–outcome effect also suggests ways to improve the quality of critical care independent of regionalization. If we are able to identify care

processes common to high-volume hospitals, it may be possible to improve outcomes by routinely exporting these processes to small community hospitals. High-volume hospitals may be more likely to utilize an a multidisciplinary care model led by trained intensivists (Young and Birkmeyer 2000). Greater nursing intensity and technological innovation may be associated with outcome in the ICU and may be correlated with higher hospital volume (Bastos et al. 1996; Tarnow-Mordi et al. 2000). High-volume hospitals may also be more likely to adopt evidence-based care processes like lung protective ventilation for acute lung injury, activated protein C for severe sepsis, and protocolized care for sedation and ventilator weaning (Ely et al. 1996; Brook et al. 1999; The Acute Respiratory Distress Syndrome Network 2000; Bernard et al. 2001). Future research should be focused on uncovering the care practices common to high-volume hospitals and investigating ways to increase use of these practices in community hospitals.

Our work has several limitations. As an observational study we cannot prove causation between volume and outcome, although our instrumental variable analysis strongly argues against patient-level confounding. Unmeasured structural characteristics may still explain part of the observed relationship—identifying these potential characteristics is in important future research direction. Of note, we observed a significant relationship even after controlling for indirect measures of technological capability and the academic status of the hospital. MediQual Atlas clinical risk adjustment was not available for all patients. The MediQual score could be missing for a number of reasons, mostly due to incomplete reporting of clinical data from the hospitals. As a consequence, the analyses using administrative risk adjustment and clinical risk adjustment were performed on slightly different groups of patients. We considered limiting the entire analysis to only patients with a MediQual score but felt this could introduce another unnecessary source of potential bias. The MediQual score was missing on a minority of patients, making it unlikely that dropping those patients in a single analysis created significant bias. Another limitation is that the MediQual Atlas score has not been validated in the ICU to the extent of other ICU-specific severity of illness measures such as the Acute Physiology and Chronic Health Evaluation (APACHE) or the Simplified Acute Physiology Score (Le Gall, Lemeshow, and Saulnier 1993; Zimmerman et al. 2006). The instrumental variable analysis produced different results than the clinical risk adjustment analysis, suggesting that MediQual Atlas did not account for all of the variation in severity of illness between volume groups. Although existing comparisons suggest that APACHE and MediQual Atlas perform similarly in hospitalized patients, this analysis

indicates that MediQual Atlas may not perform adequately in the ICU (Iezzoni 1997). Finally, our instrumental variable analysis applies only to a marginal population of ICU patients with potential to be admitted to both large- and small-volume hospitals (Harris and Remler 1998). These results would not apply to mechanically ventilated patients transferred to a large tertiary care medical center or patients requiring highly specialized ICU care available at only large volume hospitals.

CONCLUSIONS

Higher hospital volume is independently and potentially causally associated with improved outcomes for nonsurgical patients undergoing mechanical ventilation. This finding is sensitive to risk-adjustment methodology—a positive association was observed using clinical risk adjustment and an instrumental variable approach, but not with administrative risk adjustment alone. Our data resolve discordant results found in previous studies and show that instrumental variable analyses can be used to account for unmeasured confounding and selection bias in ICU outcomes research. The existence of a volume–outcome relationship in mechanical ventilation supports calls for the regionalization of critical care. Additionally, efforts to increase use of evidence-based care practices in small community hospitals may represent an opportunity to improve outcomes for critically ill patients.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: Financial support for this work was provided by a research grant from the Leonard Davis Institute of Health Economics. Dr. Kahn and Dr. Iwashyna are supported by career development awards from the National Institutes of Health (HL082650/HL091249).

Disclosures: None.

Disclaimers: The following statement is provided and required by the PHC4: PHC4 has provided these data in an effort to further PHC4's mission of educating the public and containing health care costs in Pennsylvania. PHC4, its agents and staff, have made no representation, guarantee, or warranty, expressed or implied, that the data—financial, patient, payor, and physician-specific information—are error free, or that the use of the data will avoid differences of opinion or interpretation, or disputes with those who

use published reports or purchased data. PHC4, its agents and staff, will bear no responsibility or liability for the results of the analysis, or consequences of its use.

REFERENCES

The Acute Respiratory Distress Syndrome Network. 2000. "Ventilation with Lower Tidal Volumes as Compared with Traditional Tidal Volumes for Acute Lung Injury and the Acute Respiratory Distress Syndrome." *New England Journal of Medicine* 342 (18): 1301–8.

Agresti, A. 2002. *Categorical Data Analysis.* New York: Wiley.

Angus, D. C., and N. Black. 2004. "Improving Care of the Critically Ill: Institutional and Health-Care System Approaches." *Lancet* 363 (9417): 1314–20.

Barnato, A. E., S. Alexander, W. T. Linde-Zwirble, and D. C. Angus. 2008. "Racial Variation in the Incidence, Care, and Outcomes of Severe Sepsis." *American Journal of Respiratory and Critical Care Medicine* 177: 279–84.

Barnato, A. E., J. M. Kahn, G. D. Rubenfeld, K. McCauley, D. Fontaine, J. J. Frassica, R. Hubmayr, J. Jacobi, R. G. Brower, D. Chalfin, W. Sibbald, D. A. Asch, M. Kelley, and D. C. Angus. 2007. "Prioritizing the Organization and Management of Intensive Care Services in the United States: The PrOMIS Conference." *Critical Care Medicine* 35 (4): 1003–11.

Barnato, A. E., F. L. Lucas, D. Staiger, D. E. Wennberg, and A. Chandra. 2005. "Hospital-Level Racial Disparities in Acute Myocardial Infarction Treatment and Outcomes." *Medical Care* 43 (4): 308–19.

Bastos, P. G., W. A. Knaus, J. E. Zimmerman, A. Magalhaes Jr., X. Sun, and D. P. Wagner. 1996. "The Importance of Technology for Achieving Superior Outcomes from Intensive Care. Brazil APACHE III Study Group." *Intensive Care Medicine* 22 (7): 664–9.

Beck, C. A., J. Penrod, T. W. Gyorkos, S. Shapiro, and L. Pilote. 2003. "Does Aggressive Care Following Acute Myocardial Infarction Reduce Mortality? Analysis with Instrumental Variables to Compare Effectiveness in Canadian and United States Patient Populations." *Health Services Research* 38 (6, part 1): 1423–40.

Bernard, G. R., J. L. Vincent, P. F. Laterre, S. P. LaRosa, J. F. Dhainaut, A. Lopez-Rodriguez, J. S. Steingrub, G. E. Garber, J. D. Helterbrand, E. W. Ely, and C. J. Fisher Jr. 2001. "Efficacy and Safety of Recombinant Human Activated Protein C for Severe Sepsis." *New England Journal of Medicine* 344 (10): 699–709.

Brook, A. D., T. S. Ahrens, R. Schaiff, D. Prentice, G. Sherman, W. Shannon, and M. H. Kollef. 1999. "Effect of a Nursing-Implemented Sedation Protocol on the Duration of Mechanical Ventilation." *Critical Care Medicine* 27 (12): 2609–15.

Brooks, J. M., E. A. Chrischilles, S. D. Scott, and S. S. Chen-Hardee. 2003. "Was Breast Conserving Surgery Underutilized for Early Stage Breast Cancer? Instrumental Variables Evidence for Stage II Patients from Iowa." *Health Services Research* 38 (6, part 1): 1385–402.

Deyo, R. A., D. C. Cherkin, and M. A. Ciol. 1992. "Adapting a Clinical Comorbidity Index for Use with ICD-9-CM Administrative Databases." *Journal of Clinical Epidemiology* 45 (6): 613–9.

Dombrovskiy, V. Y., A. A. Martin, J. Sunderram, and H. L. Paz. 2007. "Occurrence and Outcomes of Sepsis: Influence of Race." *Critical Care Medicine* 35 (3): 763–8.

Durairaj, L., J. C. Torner, E. A. Chrischilles, M. S. Vaughan Sarrazin, J. Yankey, and G. E. Rosenthal. 2005. "Hospital Volume–Outcome Relationships among Medical Admissions to ICUs." *Chest* 128 (3): 1682–9.

Elixhauser, A., C. Steiner, D. R. Harris, and R. M. Coffey. 1998. "Comorbidity Measures for Use with Administrative Data." *Medical Care* 36 (1): 8–27.

Ely, E. W., A. M. Baker, D. P. Dunagan, H. L. Burke, A. C. Smith, P. T. Kelly, M. M. Johnson, R. W. Browder, D. L. Bowton, and E. F. Haponik. 1996. "Effect on the Duration of Mechanical Ventilation of Identifying Patients Capable of Breathing Spontaneously." *New England Journal of Medicine* 335 (25): 1864–9.

Esteban, A., A. Anzueto, F. Frutos, I. Alia, L. Brochard, T. E. Stewart, S. Benito, S. K. Epstein, C. Apezteguia, P. Nightingale, A. C. Arroliga, and M. J. Tobin. 2002. "Characteristics and Outcomes in Adult Patients Receiving Mechanical Ventilation: A 28-Day International Study." *Journal of the American Medical Association* 287 (3): 345–55.

Glance, L. G., Y. Li, T. M. Osler, A. Dick, and D. B. Mukamel. 2006. "Impact of Patient Volume on the Mortality Rate of Adult Intensive Care Unit Patients." *Critical Care Medicine* 34 (7): 1925–34.

Halm, E. A., C. Lee, and M. R. Chassin. 2002. "Is Volume Related to Outcome in Health Care? A Systematic Review and Methodologic Critique of the Literature." *Annals of Internal Medicine* 137 (6): 511–20.

Hannan, E. L., H. Kilburn Jr., M. L. Lindsey, and R. Lewis. 1992. "Clinical Versus Administrative Data Bases for CABG Surgery. Does It Matter?" *Medical Care* 30 (10): 892–907.

Harris, K. M., and D. K. Remler. 1998. "Who Is the Marginal Patient? Understanding Instrumental Variables Estimates of Treatment Effects." *Health Services Research* 33 (5, part 1): 1337–60.

Harris, M. I. 2001. "Racial and Ethnic Differences in Health Care Access and Health Outcomes for Adults with Type 2 Diabetes." *Diabetes Care* 24 (3): 454–9.

Hausman, J., and D. McFadden. 1984. "Specification Tests in Econometics." *Econometrica* 46: 1251–71.

Iezzoni, L. I. 1997. "The Risks of Risk Adjustment." *Journal of American Medical Association* 278 (19): 1600–7.

Jacobs, P., J. Rapoport, and D. Edbrooke. 2004. "Economies of Scale in British Intensive Care Units and Combined Intensive Care/High Dependency Units." *Intensive Care Medicine* 30 (4): 660–4.

Kahn, J. M. 2007. "Volume, Outcome, and the Organization of Intensive Care." *Critical Care* 11 (3): 129.

Kahn, J. M., C. H. Goss, P. J. Heagerty, A. A. Kramer, C. R. O'Brien, and G. D. Rubenfeld. 2006. "Hospital Volume and the Outcomes of Mechanical Ventilation." *New England Journal of Medicine* 355 (1): 41–50.

Kahn, J. M., W. T. Linde-Zwirble, H. Wunsch, A. E. Barnato, T. J. Iwashyna, M. S. Roberts, J. R. Lave, and D. C. Angus. 2008. "Potential Value of Regionalized Intensive Care for Mechanically Ventilated Medical Patients." *American Journal of Respiratory and Critical Care Medicine* 177: 285–91.

Kennedy, P. 2002. *A Guide to Econometrics*. Cambridge, MA: The MIT Press.

Konety, S. H., M. S. Vaughan Sarrazin, and G. E. Rosenthal. 2005. "Patient and Hospital Differences Underlying Racial Variation in Outcomes after Coronary Artery Bypass Graft Surgery." *Circulation* 111 (10): 1210–6.

Landon, B. E., S. L. Normand, A. Lessler, A. J. O'Malley, S. Schmaltz, J. M. Loeb, and B. J. McNeil. 2006. "Quality of Care for the Treatment of Acute Medical Conditions in US Hospitals." *Archives of Internal Medicine* 166 (22): 2511–7.

Le Gall, J. R., S. Lemeshow, and F. Saulnier. 1993. "A New Simplified Acute Physiology Score (SAPS II) Based on a European/North American Multicenter Study." *Journal of the American Medical Association* 270 (24): 2957–63.

Luft, H. S., S. S. Hunt, and S. C. Maerki. 1987. "The Volume–Outcome Relationship: Practice-Makes-Perfect or Selective-Referral Patterns?" *Health Services Research* 22 (2): 157–82.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, England: Cambridge University Press.

McClellan, M., B. J. McNeil, and J. P. Newhouse. 1994. "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? Analysis Using Instrumental Variables." *Journal of the American Medical Association* 272 (11): 859–66.

Needham, D. M., S. E. Bronskill, D. M. Rothwell, W. J. Sibbald, P. J. Pronovost, A. Laupacis, and T. A. Stukel. 2006. "Hospital Volume and Mortality for Mechanical Ventilation of Medical and Surgical Patients: A Population-Based Analysis Using Administrative Data." *Critical Care Medicine* 34 (9): 2349–54.

Peelen, L., N. F. De Keizer, N. Peek, G. J. Scheffer, P. H. Van der Voort, and E. De Jonge. 2007. "The Influence of Volume and ICU Organization on Hospital Mortality in Patients Admitted with Severe Sepsis: A Retrospective Multicenter Cohort Study." *Critical Care* 11 (2): R40.

Petersen, L. A., S. M. Wright, E. D. Peterson, and J. Daley. 2002. "Impact of Race on Cardiac Care and Outcomes in Veterans with Acute Myocardial Infarction." *Medical Care* 40 (1, suppl): I86–96.

Pine, M., H. S. Jordan, A. Elixhauser, D. E. Fry, D. C. Hoaglin, B. Jones, R. Meimban, D. Warner, and J. Gonzales. 2007. "Enhancement of Claims Data to Improve Risk Adjustment of Hospital Mortality." *Journal of the American Medical Association* 297 (1): 71–6.

Rivers, D., and Q. H. Vuong. 1988. "Limited Information Estimators and Exogenity Tests for Simultaneous Probit Models." *Journal of Econometrics* 39: 347–66.

Stukel, T. A., E. S. Fisher, D. E. Wennberg, D. A. Alter, D. J. Gottlieb, and M. J. Vermeulen. 2007. "Analysis of Observational Studies in the Presence of Treatment Selection Bias: Effects of Invasive Cardiac Management on AMI Survival Using Propensity Score and Instrumental Variable Methods." *Journal of the American Medical Association* 297 (3): 278–85.

Tarnow-Mordi, W. O., C. Hau, A. Warden, and A. J. Shearer. 2000. "Hospital Mortality in Relation to Staff Workload: A 4-Year Study in an Adult Intensive-Care Unit." *Lancet* 356 (9225): 185–9.

Tsai, A. C., M. Votruba, J. F. Bridges, and R. D. Cebul. 2006. "Overcoming Bias in Estimating the Volume–Outcome Relationship." *Health Services Research* 41 (1): 252–64.

Williams, J. F., J. E. Zimmerman, D. P. Wagner, M. Hawkins, and W. A. Knaus. 1995. "African-American and White Patients Admitted to the Intensive Care Unit: Is There a Difference in Therapy and Outcome?" *Critical Care Medicine* 23 (4): 626–36.

Young, M. P., and J. D. Birkmeyer. 2000. "Potential Reduction in Mortality Rates Using an Intensivist Model to Manage Intensive Care Units." *Effective Clinical Practice* 3 (6): 284–9.

Zimmerman, J. E., A. A. Kramer, D. S. McNair, and F. M. Malila. 2006. "Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital Mortality Assessment for Today's Critically Ill Patients." *Critical Care Medicine* 34 (5): 1297–310.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.