

HOSPITAL MEDICINE

Did Duty Hour Reform Lead to Better Outcomes Among the Highest Risk Patients?

Kevin G. Volpp, MD, PhD^{1,2,3,4}, Amy K. Rosen, PhD^{5,6}, Paul R. Rosenbaum, PhD⁷, Patrick S. Romano, MD, MPH⁸, Kamal M.F. Itani, MD⁹, Lisa Bellini, MD², Orit Even-Shoshan, MS^{10,13}, Liyi Cen, MS², Yanli Wang, MS¹⁰, Michael J. Halenar, BA^{1,2}, and Jeffrey H. Silber, MD, PhD^{3,10,11,12,13}

¹Center for Health Equity Research and Promotion, Veteran's Administration Hospital, Philadelphia, PA, USA; ²Department of Medicine, The University of Pennsylvania School of Medicine, Philadelphia, PA, USA; ³Department of Health Care Management, The Wharton School, The University of Pennsylvania, Philadelphia, PA, USA; ⁴Center for Health Incentives, The Leonard Davis Institute of Health Economics, The University of Pennsylvania, Philadelphia, PA, USA; ⁵Department of Health Policy and Management, Boston University School of Public Health, Boston, MA, USA; ⁶Center for Health Quality, Outcomes and Economic Research, Veteran's Administration Hospital, Bedford, MA, USA; ⁷Department of Statistics, The Wharton School, The University of Pennsylvania, Philadelphia, PA, USA; ⁸Division of General Medicine and Center for Healthcare Policy and Research, University of California Davis School of Medicine, Davis, CA, USA; ⁹Department of Surgery, VA Boston Health Care System and Boston University, Boston, MA, USA; ¹⁰Center for Outcomes Research, The Children's Hospital of Philadelphia, Philadelphia, PA, USA; ¹¹The Department of Pediatrics, The University of Pennsylvania School of Medicine, Philadelphia, PA, USA; ¹²The Department of Anesthesiology and Critical Care, The University of Pennsylvania School of Medicine, Philadelphia, PA, USA; ¹³The Leonard Davis Institute of Health Economics, The University of Pennsylvania, Philadelphia, PA, USA.

BACKGROUND: Earlier work demonstrated that ACGME duty hour reform did not adversely affect mortality, with slight improvement noted among specific subgroups.

OBJECTIVE: To determine whether resident duty hour reform differentially affected the mortality risk of high severity patients or patients who experienced post-operative complications (failure-to-rescue).

DESIGN: Observational study using interrupted time series analysis with data from July 1, 2000 - June 30, 2005. Fixed effects logistic regression was used to examine the change in the odds of mortality or failure-to-rescue (FTR) in more versus less teaching-intensive hospitals before and after duty hour reform.

PARTICIPANTS: All unique Medicare patients (n=8,529,595) admitted to short-term acute care non-federal hospitals and all unique VA patients (n=318,636 patients) with principal diagnoses of acute myocardial infarction, congestive heart failure, gastrointestinal bleeding, stroke or a DRG classification of general, orthopedic or vascular surgery.

MEASUREMENTS AND MAIN RESULTS: We measured mortality within 30 days of hospital admission and FTR, measured by death among patients who experienced a surgical complication. The odds of mortality and FTR generally changed at similar rates for higher and lower risk patients in more vs. less teaching intensive hospitals. For example, comparing the mortality risk for the 10% of Medicare patients with highest risk to the other 90% of patients in post-reform year 1 for combined medical an OR of 1.01 [95% CI 0.90, 1.13], for combined surgical an OR of 0.91 [95% CI 0.80, 1.04],

and for FTR an OR of 0.94 [95% CI 0.80, 1.09]. Findings were similar in year 2 for both Medicare and VA. The two exceptions were a relative increase in mortality for the highest risk medical (OR 1.63 [95% CI 1.08, 2.46]) and a relative decrease in the high risk surgical patients within VA in post-reform year 1 (OR 0.52 [95% CI 0.29, 0.96]).

CONCLUSIONS: ACGME duty hour reform was not associated with any consistent improvements or worsening in mortality or failure-to-rescue rates for high risk medical or surgical patients.

KEY WORDS: medical errors internship and residency; education, medical, graduate; personnel staffing and scheduling; continuity of patient care.

J Gen Intern Med 24(10):1149-55

DOI: 10.1007/s11606-009-1011-z

© Society of General Internal Medicine 2009

INTRODUCTION

Recent work has demonstrated that the ACGME duty hour regulations for physicians in training did not lead to increased mortality¹ and may have been associated with improved mortality for subsets of patients.^{2,3} However, these studies reported the average impacts across the entire population, and any observed impact on patient outcomes might only manifest in higher severity patients. High severity patients have less physiologic reserve and may be more susceptible to either benefits or harms from changes in the way care is delivered in teaching hospitals.

At teaching hospitals, residents play an important role in recognizing complications and implementing corrective action in high-risk patients. If duty hour reform mitigated resident fatigue, then residents' ability to recognize and react to complications may have improved, thereby improving mortality among high-risk patients and among surgical patients who

Received November 10, 2008

Revised March 25, 2009

Accepted April 6, 2009

Published online May 20, 2009

had already experienced complications (failure-to-rescue or FTR rates). On the other hand, duty hour reform has resulted in fewer residents on duty at any one time, the work intensity of residents may have increased, and continuity of care might have worsened, leading to higher rates of complications, FTR, and mortality in high risk patients.⁴⁻⁷

The objective of this study is to determine how duty hour reform affected high risk hospitalized patients by examining changes in outcomes in two groups of high severity patients: (1) those with the highest mortality risk (top 10% or 25%) based on characteristics on admission; (2) surgical patients who suffered a complication. Among surgical patients who suffered a complication, we examined the FTR rate, a quality of care measure that reflects the capacity of a hospital to provide the care needed to prevent death after a patient suffers a complication.⁸⁻¹¹

METHODS

Study Sample

The **Resident Duty Hours Study** utilized all patients admitted to Veterans Health Administration (VA) medical centers and Medicare patients admitted to short-term general non-federal acute-care hospitals from July 1, 2000 to June 30, 2005 with a principal diagnosis of acute myocardial infarction (AMI), congestive heart failure (CHF), gastrointestinal (GI) bleeding, or stroke, or with a DRG classification of general, orthopedic, or vascular surgery. Details of the sample have been previously reported.^{1,2} In short, for the mortality analyses we examined data on 318,636 patients from 131 VA hospitals and 8,529,595 Medicare patients from 3,321 hospitals. The failure-to-rescue analyses were done among surgical patients with principal procedure/DRG classification of general, orthopedic or vascular surgery. The initial surgical sample included 224,485 VA patients from 142 hospitals and 6,610,766 Medicare surgical patients from 5,736 acute care hospitals within 50 states. Exclusions resulted in a total sample for analysis of 67,830 VA patients from 129 VA hospitals and 4,658,594 Medicare patients from 3,302 hospitals.

Risk Adjustment and Hospital Control Measures

The risk-adjustment approach used was developed by Elixhauser et al.¹² It uses 27 comorbidities (excluding fluid and electrolyte disorders and coagulopathy, diagnoses which may indicate complications rather than comorbidities)^{13,14} and has been shown to achieve better discrimination than alternative approaches.^{15,16} This approach was augmented with adjustments for age and sex. We also adjusted for DRGs for surgical patients, grouping paired DRGs with and without complications or comorbidities into one aggregated DRG to avoid adjusting for potentially iatrogenic events. We performed a 180-day "lookback" including data from prior hospitalizations, to obtain more comprehensive information on comorbidities than available using the index admission alone.^{11,17}

Risk Score Calculation

Risk scores were derived as follows: (1) using data from 1999-2000 (the year prior to the study sample), logistic regression

was used to obtain coefficients for each patient-level variable including the 27 comorbidities described above, age, gender, transfer in status, and a set of DRG and procedure combinations based on Haberman residuals to create homogenous groups with respect to mortality;^{11,18-20} (2) risk scores were then calculated by applying the coefficients from step (1) to the study sample. The risk score for a patient is an estimate of the probability of death within 30 days, so a high risk patient is one with a high estimated probability of death. These risk scores (for example, VA medical: 5th percentile 1.5%, 95th percentile 21.3%; Medicare medical: 5th percentile 3.3%, 95th percentile 28.6%) were then used to stratify the sample for analysis in four groups: low risk (below 25th percentile), medium risk (25-75th percentile), high risk (75th-90th percentile), and very high risk (>90th percentile).

Teaching Intensity: The Resident-to-Bed Ratio

The number of residents per hospital was obtained from the Medicare Cost Reports from the Centers for Medicare and Medicaid Services and from the VA Office of Academic Affiliations. The resident/bed ratio (defined as the ratio of (interns + residents)/average operating beds (RB ratio)^{1,2,21} has been used a measure of teaching intensity in previous studies,²²⁻²⁴ differentiating hospitals as either major, minor, or non-teaching. The RB ratio measure is an ideal measure to study the influence of the change in duty hours; the greater the number of residents per bed, the more likely we are to observe a change if a change in outcomes did occur.^{1,2}

We used the resident/bed ratio as a continuous variable, which provided more power for assessing associations with duty hour rule implementation than dividing hospitals into arbitrary categories.^{25,26} We held the resident/bed ratio fixed using the level in pre-reform year 1 so that any potential behavioral response to the reforms by hospitals (such as changing the number of residents) would not confound estimation of the net effects of duty hour reform.

Complications and Failure-to-Rescue

Failure-to-Rescue (FTR) is defined as the probability of death following a complication.⁸⁻¹¹ FTR focuses on how a complication is managed in order to avoid a death, where the probability of death is equal to the probability of dying given a complication multiplied by the probability of developing a complication. Ideally, all complications for all patients should be recorded. However, there may be deaths that are not preceded by a documented complication. This may be because deaths occurred outside the hospital and the complications were undocumented or the hospital did not record some complications prior to death. In our FTR analyses we assume all deaths are preceded by at least one complication, as all patients who die must have had, for example, a ventricular arrhythmia or asystole. The complications utilized in our FTR analyses have been described previously¹¹ and represent a wide range of complications.

Statistical Analysis

We used a multiple time series research design,²⁷ also known as "difference-in-differences", to examine whether the implementation of duty hour reform was associated with a change in

the underlying trend in patient outcomes for high risk patients in teaching hospitals. This approach reduces certain potential biases from unmeasured variables.^{28,29} The multiple time series research design compares each hospital to itself, before and after reform, contrasting the changes in hospitals with more residents to the changes in hospitals with fewer or no residents, making adjustments for observed differences in patient risk factors. It also adjusts for changes in outcomes over time (trends) that were common to all hospitals. In this design, a time invariant difference between hospitals cannot be confused with an effect of the reform. Similarly, a trend over time that affects all hospitals in the same way cannot be confused with an effect of the reform. Finally, a change in the patient mix, to the extent that it is reflected in measured patient characteristics, cannot be confused with an effect of the reform. The advantages of this design have been previously reported.^{1,2}

Less teaching-intensive hospitals, including all non-teaching hospitals, served as the primary control group for more teaching-intensive hospitals because they were less affected by the duty hour reform but were subject to the same technological quality improvement imperatives, changes in market conditions, and Medicare-specific initiatives such as 'pay for performance' or VA quality improvement efforts. In addition, they are geographically diverse with large patient populations, and similar patient discharge data are available. Data from July 1, 2000-June 30, 2003 were used as the pre-reform period, with data from July 1, 2003-June 30, 2005 serving as the post-reform period.

The dependent variable was death within 30 days of hospital admission, using logistic regression to adjust for patient comorbidities, secular trends common to all patients (for example, due to general changes in technology), and hospital site where treated. The change in outcomes associated with duty hour reform in highest risk patients was measured using the coefficients of resident/bed ratio interacted with dummy variables that indicate having a risk score in the top 10% of all patients (or top 25% of all patients) and dummy variables indicating post-reform year 1 (July 1, 2003-June 30, 2004) and post-reform year 2 (July 1, 2004-June 30, 2005). These triple interaction coefficients, presented as odds ratios, measure the degree to which mortality changed among the highest risk patients in more vs less teaching-intensive hospitals compared to lower risk patients after adjusting for cross-sectional differences in hospital quality and general improvements in care. They were measured for each year separately because of the possibility of either delayed beneficial effects or early harmful effects. Conditions were assessed for all medical patients as a group and all surgical patients as a group. FTR analyses were done on surgical patients only, since the coding of complications is more likely to be accurate for surgical than medical patients.³⁰⁻³²

In the models, baseline mortality levels for high risk patients were allowed to differ between more and less teaching-intensive hospitals and were assumed to have a common time trend until implementation of the duty hour rules, after which the teaching hospital trend was allowed to diverge. To assess whether underlying trends in risk-adjusted mortality were similar in higher and lower teaching intensity hospitals during the 3 years prior to the ACGME duty hour reform, a "test of controls" was performed. Parameters were added to the model for interactions between being a high severity patient, the resident/bed ratio, and indicators for pre-reform year 2 and pre-reform year 1. A Wald test was used to determine whether

these interactions were equal to 0. A statistically significant test of controls suggests that high and low risk patients had diverging trends in more vs. less teaching-intensive hospitals in the 3 years pre-reform that could not have been caused by the reform. When such diverging trends were found by the test of controls, post hoc analyses were conducted in which post-reform results were compared to the pre-reform year 1 (July 1, 2002-June 30, 2003) as a baseline rather than using data from the entire 3-year pre-reform period.

We tested the stability of the results by (1) eliminating patients admitted to hospitals in New York State, due to earlier passage of the Libby Zion laws; and (2) assessing whether the results changed when we eliminated all patients who had metastatic cancer (the distribution of severity, e.g. patients who were in the top 10% of severity, was recalculated after excluding patients with a secondary diagnosis of metastatic cancer). All p-values are either 2-tailed or, for chi-squared tests, multi-tailed. A p-value of <0.05 was considered statistically significant. All analyses were conducted with SAS 9.1 (SAS Institute, Inc., Cary, NC).

RESULTS

Patients in the higher risk groups were older and generally had a higher mean number of comorbidities (Table 1 for medical patients, Table 2 for surgical patients). In particular, higher risk patients had a much higher prevalence of comorbidities that were strongly associated with mortality. For example, the proportion of medical patients within the very high risk group (top 10% of risk) who had metastatic cancer was 12.1% within Medicare and 13.6% within VA compared to 0.0% among low risk patients (bottom 25% of risk) in both populations (Table 1). Expressed differently, 47% of Medicare and 83% of VA patients with metastatic cancer were in the very high risk group (not shown in table). Similar patterns were seen for liver disease, renal failure, and congestive heart failure. A diagnosis of hypertension on administrative data tends to be protective,³³ and a higher percentage of patients with hypertension was observed in the lower risk groups. Patterns were similar among surgical patients (Table 2). The distribution of patients of differing severity was similar in hospitals of differing teaching intensity among both VA and Medicare medical and surgical patients (Tables 1 and 2).

Examination of unadjusted trends in mortality for very high risk Medicare patients indicated that changes in mortality rates for high risk patients did not vary with teaching intensity (Fig. 1). Among VA medical and surgical patients, the highest severity patients within the most teaching-intensive hospitals experienced a relative increase in mortality in post-reform year 1 compared to medical patients in less teaching-intensive hospitals. A relative decrease in mortality was seen in post-reform year 1 for the highest severity VA surgical patients in more teaching-intensive hospitals.

In examining the adjusted results among Medicare patients, we found that in the combined medical group the differential in mortality in more vs. less teaching-intensive hospitals (the 'differential') changed at similar rates among patients in the highest 10% of risk and the other 90% of patients in both post-reform years 1 (OR 1.01 [95% CI 0.90, 1.13]) and 2 (OR 0.90 [95% CI 0.80, 1.02]) (Table 3). This group was the one group in which the test of controls indicated that pre-reform trends

Table 1. Characteristics of the Study Population - Combined Medical

	Low Risk (<25%)	Medium Risk (25–75%)	High Risk (75–90%)	Very High Risk (≥90%)
Medicare				
Number patients	966,108	1,931,170	580,012	386,178
Mean age	75.08	77.74	81.91	82.95
% male	42.3	46.2	46.2	49.9
Mean number comorbidities	2.13	1.98	1.96	2.47
Metastatic Cancer (%)	0.02	0.74	2.24	12.12
Liver disease (%)	0.72	1.08	1.39	2.71
Renal Failure (%)	1.45	5.41	9.82	22.98
Hypertension (%)	13.13	11.20	9.59	9.03
CHF (%)	4.91	12.27	17.81	43.88
Teaching Intensity ¹				
RB ratio=0	26.0	49.4	14.8	9.9
RB ratio >0–0.25	24.1	50.7	15.2	10.0
RB ratio 0.26–0.60	23.8	50.3	15.3	10.0
RB ratio >0.60	22.9	51.4	15.5	10.2
VA				
Number patients	36,470	72,935	21,910	14,590
Mean age	58.0	69.1	75.4	76.9
% male	97.0	98.5	98.8	99.0
Mean number comorbidities	2.0	2.1	2.4	3.2
Metastatic Cancer (%)	0.0	0.2	1.3	13.6
Liver disease (%)	0.9	4.3	6.1	10.6
Renal Failure (%)	0.8	3.5	11.4	29.0
Hypertension (%)	73.9	62.2	48.3	42.3
CHF (%)	1.7	6.6	15.3	29.3
Teaching Intensity ¹				
RB ratio=0	20.1	50.7	17.8	11.5
RB ratio >0–0.25	24.0	49.8	15.8	10.4
RB ratio 0.26–0.60	25.5	50.2	14.7	9.6
RB ratio >0.60	25.6	49.8	14.6	10.0

¹ Teaching intensity is presented as a row percentage (specifically, the percentage of patients in each category of teaching hospitals who are within that risk stratum). All other percentages given are column percentages (prevalence of each comorbidity among patients in each risk stratum).

were different in more vs. less teaching-intensive hospitals so pre-reform year 1 was used as the baseline, and results were qualitatively similar. Findings were also similar for patients in the top 25% of risk compared to all others. Among Medicare patients in the combined surgical group, the differential in mortality changed at similar rates among patients in the highest 10% of risk and the other 90% of patients in both post-reform years 1 (OR 0.91 [95% CI 0.80, 1.04]) and 2 (OR 1.01 [95% CI 0.88, 1.15]) (Table 3). Findings were similar for patients in the top 25% of risk compared to all others. These findings should be interpreted in the context of our previous findings of no significant relative changes in mortality for medical or surgical patients in more vs. less teaching-intensive hospitals in either post-reform year 1 or 2.¹

Among Medicare surgical patients in the top 10% of risk, FTR rates changed at similar rates compared to patients in the bottom 90% of risk in more vs. less teaching-intensive hospitals (OR 0.94 [95% CI 0.80, 1.09]). Findings were similar for patients in the top 25% of risk compared to all others. Among all Medicare surgical patients, the FTR rate overall (not previously reported or shown in table) changed at similar rates in more vs. less teaching-intensive hospitals in both post-reform year 1 (OR 1.04 [95% CI 0.91, 1.11]) and post-reform year 2 (OR 1.00 [95% CI 0.94, 1.07]).

In stability analyses among Medicare patients in the highest 10% of risk, we found that for combined medical patients excluding patients admitted to hospitals in NY produced qualitatively similar results in post-reform year 1. However, the odds of mortality for patients in the top 10% of risk decreased in more vs. less teaching intensive hospitals to a greater degree than among the other 90% of patients (OR 0.87, 95% CI 0.77, 0.99) in post-reform year 2. Excluding patients who had metastatic cancer produced qualitatively similar results as the primary analyses. Exclusion of either patients admitted to hospitals in NY or patients with metastatic cancer produced qualitatively similar results for both the combined surgical mortality and the FTR analyses.

Among VA patients, the test of controls showed no evidence that pre-reform trends were different in more vs. less teaching-intensive hospitals so the entire pre-reform period was used as the baseline for all analyses. Adjusted analyses indicated that the odds of mortality for medical patients in the highest 10% of severity increased to a greater degree than among the other

Table 2. Characteristics of the Study Population - Combined Surgical

	Low Risk (<25%)	Medium Risk (25–75%)	High Risk (75–90%)	Very High Risk (≥90%)
Medicare				
Number patients	1,166,148	2,333,318	700,392	466,269
Mean age	73.09	75.84	80.18	81.14
% male	29.49	36.94	47.97	53.37
Mean number comorbidities	1.44	1.67	2.14	2.68
Metastatic Cancer (%)	0.26	1.78	9.53	21.15
Liver disease (%)	0.19	0.68	1.88	4.87
Renal Failure (%)	0.28	1.16	5.33	18.03
Hypertension (%)	14.01	11.99	8.43	7.50
CHF (%)	1.82	5.34	22.25	53.83
Complications rate	22.55	40.95	60.62	75.54
FTR rate	1.85	5.88	12.89	22.13
Teaching Intensity ¹				
RB ratio=0	24.4	50.7	15.2	9.8
RB ratio >0–0.25	25.7	49.7	14.7	9.9
RB ratio >0.25–0.60	25.1	48.7	15.2	11.1
RB ratio >0.60	25.9	48.2	15.1	10.7
VA				
Number patients	43,695	87,390	26,217	17,478
Mean age	51.3	63.8	69.9	73.5
% male	89.2	97.2	98.1	98.9
Mean number comorbidities	0.9	1.4	2.2	3.0
Metastatic Cancer (%)	0.2	1.8	9.5	18.3
Liver disease (%)	0.6	3.0	4.7	11.2
Renal Failure (%)	0.2	1.8	6.4	15.9
Hypertension (%)	36.5	44.7	50.1	47.3
CHF (%)	0.5	3.9	9.6	22.9
Complications rate	17.0	37.7	58.2	69.7
FTR rate	1.7	3.6	10.5	19.3
Teaching Intensity ¹				
RB ratio=0	19.0	52.4	17.0	11.6
RB ratio >0–0.25	22.1	52.4	15.7	9.8
RB ratio 0.26–0.60	24.9	49.4	15.6	10.1
RB ratio >0.60	26.1	49.9	14.1	10.0

¹ Teaching intensity is presented as a row percentage (specifically, the percentage of patients in each category of teaching hospitals who are within that risk stratum). All other percentages given are column percentages (prevalence of each comorbidity among patients in each risk stratum).

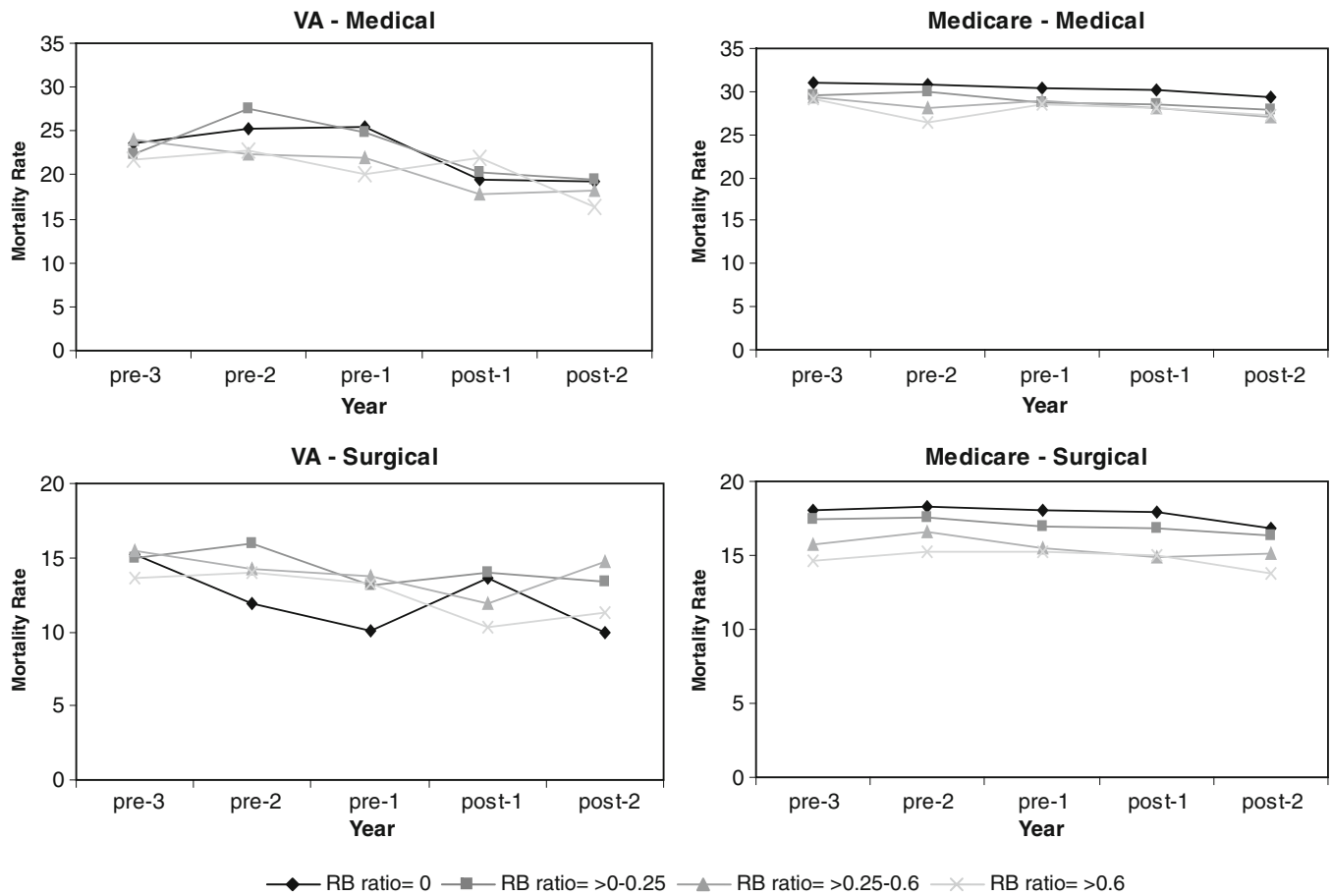


Figure 1. Changes over time in unadjusted mortality for very high severity patients in hospitals of different teaching intensity.

90% of patients in more vs. less teaching-intensive hospitals in the first year post-reform (OR 1.63, [95% CI 1.08, 2.46]), with the relative odds of mortality being non-significantly higher by year 2 (OR 1.35, [95% CI 0.88, 2.07]). A qualitatively similar pattern was observed among patients in the highest 25% of risk compared to the lowest 75%. Among VA surgical patients, a relative decrease in the odds of mortality was not observed for patients in the top 10% of risk compared to lower risk patients in more vs. less teaching-intensive hospitals (OR 0.68 [95% CI 0.39, 1.20]). Patients in the highest 25% of risk experienced lower odds of mortality relative to other patients in more vs. less teaching-intensive hospitals (OR 0.52 [95% CI 0.29, 0.96]). These findings should be viewed in the context of our previously reported findings of no overall change in mortality in more vs. less teaching-intensive VA hospitals in post-reform year 1 among medical or surgical patients, but significant relative improvements in mortality for medical (but not surgical) patients in more teaching-intensive VA hospitals in the second year post-reform only.²

FTR rates among VA patients in the highest 10% of risk changed at rates similar to that of other patients in more vs. less teaching-intensive hospitals in both post-reform year 1 (OR 0.67 [95% CI 0.35, 1.30]) and year 2 (OR 0.64 [95% CI 0.33, 1.24]). Similar patterns were observed in comparing changes in the odds of FTR among the highest 25% risk patients vs. all others. The overall FTR rates (not previously

reported and not shown in table) changed at similar rates in more vs. less teaching-intensive hospitals in post-reform year 1 (OR 0.87 [95% CI 0.65, 1.15]) and year 2 (OR 0.94 [95% CI 0.70, 1.26]).

Excluding patients admitted to hospitals in New York State or patients who had metastatic cancer produced qualitatively similar results in both the mortality and FTR analyses.

DISCUSSION

In a national study of both patients treated by Medicare and Veterans Administration hospitals, our findings suggest that ACGME duty hour reform had no systematic impact on the outcomes of high as compared to low risk patients. These findings are important in building on previous work that suggested no systematic changes, on average, in mortality.¹⁻³ These data could be viewed as evidence that even among the patients in the highest ten percent of risk, duty hour reform neither worsened nor improved their mortality risk nor their likelihood of avoiding death following a surgical complication.

Previous work by Shetty and Bhattacharya suggested that medical patients who were older as well as patients with certain comorbidities experienced larger improvements in outcomes than other patients.³ However, this previous work had some important limitations including use of in-hospital mor-

Table 3. Odds of Mortality and Failure-to-Rescue Post duty hour Reform in More vs. Less Teaching-intensive Hospitals

	Resident/bed ratio × post-reform year 1 * high/very high risk ¹		Resident/bed ratio × post-reform year 2 * high/very high risk ¹	
	Odds ratio† (95% CI)	P-value	Odds ratio† (95% CI)	P-value
Medicare				
<i>Combined medical</i>				
Highest 10% (vs. bottom 90%)	1.01 (0.90, 1.13)	0.86	0.90 (0.80, 1.02)	0.09
Highest 25% (vs. bottom 75%)	0.99 (0.98, 1.08)	0.81	0.94 (0.85, 1.03)	0.17
<i>Combined surgical</i>				
Highest 10% (vs. bottom 90%)	0.91 (0.80, 1.04)	0.18	1.01 (0.88, 1.15)	0.88
Highest 25% (vs. bottom 75%)	0.98 (0.86, 1.12)	0.76	1.09 (0.95, 1.24)	0.21
<i>FTR rates</i>				
Highest 10% (vs. bottom 90%)	0.94 (0.80, 1.09)	0.40	1.01 (0.86, 1.18)	0.92
Highest 25% (vs. bottom 75%)	0.90 (0.79, 1.02)	0.10	1.00 (0.88, 1.14)	0.98
VA				
<i>Combined medical</i>				
Highest 10% (vs. bottom 90%)	1.63 (1.08, 2.46)	0.02	1.35 (0.88, 2.07)	0.17
Highest 25% (vs. bottom 75%)	1.44 (1.01, 2.05)	0.045	0.99 (0.67, 1.43)	0.93
<i>Combined surgical</i>				
Highest 10% (vs. bottom 90%)	0.68 (0.39, 1.20)	0.19	0.80 (0.45, 1.43)	0.45
Highest 25% (vs. bottom 75%)	0.52 (0.29, 0.96)	0.04	1.13 (0.59, 2.17)	0.71
<i>FTR rates</i>				
Highest 10% (vs. bottom 90%)	0.67 (0.35, 1.30)	p=0.23	0.64 (0.33, 1.24)	p=0.19
Highest 25% (vs. bottom 75%)	0.86 (0.49, 1.51)	p=0.60	0.82 (0.46, 1.48)	p=0.51

*The interaction terms (high severity × resident/bed ratio × post-reform year 1) and (high severity × resident/bed ratio × post-reform year 2) measure whether there is any relative change in mortality or the failure-to-rescue rate in more versus less teaching-intensive hospitals for high severity patients compared to low severity patients. There are two models: one which compares changes in outcomes for patients in the highest 10% of severity with all other patients and one which compares changes in outcomes for patients in the highest 25% of severity with all other patients. Models are also adjusted for age, sex, comorbidities, common time trends, and hospital site where treated.

†Units for odds ratios compare hospitals with resident/bed ratio=1 to hospitals with resident/bed ratio=0 (non-teaching).

tality as the primary outcome, creating the potential for bias from differential changes in length of stay among teaching and non-teaching hospitals; use of data that could not differentiate between two admissions from the same patient or two admissions from different patients; and no examination of the synergistic impact of combinations of comorbidities and other mortality risk factors such as age on mortality outcomes. Nonetheless, this work coupled with our current study highlights that higher risk patients did not appear to have been harmed by duty hour reform and may have derived some benefits.

Current deliberations about further modifications to the duty hour regulations to improve patient safety³⁴ must weigh the evidence on the lack of beneficial or harmful effects on mortality of the current ACGME duty hour standards against evidence that residents have a higher quality of life under the new system.^{35,36} There is still little evidence about how the duty hour reform affected outcomes other than mortality. In related work,^{37,38} we have found no consistent changes in the probability of patients experiencing prolonged length of stay (other than possibly in vascular surgery) or in the rates at which a wide range of patients experienced complications (Patient Safety Indicators or PSIs) before versus after the duty hour changes in more and less teaching-intensive hospitals. A prolonged stay or PSI event may not be as consequential as a death, but both were examined to look for more subtle effects of the change in duty hour rules that could reflect changes in either efficiency of care or the probability of developing complications that lead to delays in patient discharges. Studies of the impact on the quality of resident training as well as how patient outcomes would compare under the current system vs. other alternatives such as 16-hour shifts^{39,40} are needed to better inform further policy decisions about duty hour regulation.

Limitations of this study include our focus on two measures that use mortality as the outcome. Even with the size of the Medicare and VA populations, some of the confidence intervals are still somewhat broad and we cannot rule out small and possibly clinically meaningful effects, particularly in the context of FTR for VA surgical patients. Another potentially significant limitation in any observational study of this type is unmeasured confounding. However, by comparing outcomes over time within each hospital, in more versus less teaching-intensive hospitals, potential bias from unmeasured cofounders is diminished. We are also limited in our ability to do risk adjustment using administrative data, but FTR may be more sensitive to hospital characteristics and actions as compared to patient case-mix and severity than mortality rates,^{11,18} and findings using FTR as an outcome were very similar to our findings using surgical mortality.

In conclusion, the preponderance of the evidence indicates that higher severity patients were not disproportionately harmed or benefitted compared to lower risk patients following implementation of duty hour reform. Future work should assess impacts on different measures of clinical outcomes, educational outcomes, and, ideally, include direct comparisons of different approaches to duty hour regulation.

Acknowledgements: This work was supported primarily by grant VA HSR&D IIR 04.202.1 and NHLBI R01 HL082637, with additional support from National Science Foundation grant SES-0646002. The sponsors/funders have had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. Dr. Volpp had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. We also thank Yun Teng for her assistance with the

analyses. Everyone who contributed significantly to this work has been acknowledged.

Conflict of Interest: None disclosed.

Corresponding Author: Kevin G. Volpp, MD, PhD: Center for Health Equity Research and Promotion, Veteran's Administration Hospital, 1232 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, USA (e-mail: volpp70@wharton.upenn.edu).

REFERENCES

1. **Volpp KG, Rosen AK, Rosenbaum PR, et al.** Mortality among hospitalized Medicare beneficiaries in the first two years following ACGME resident duty hour reform. *JAMA*. 2007;298:975-83.
2. **Volpp KG, Rosen AK, Rosenbaum PR, et al.** Mortality among patients in VA hospitals in the first two years following ACGME resident duty hour reform. *JAMA*. 2007;298:984-92.
3. **Shetty KD, Bhattacharya J.** Changes in hospital mortality associated with residency work-hour regulations. *Ann Intern Med*. 2007;147(2):73-80.
4. **Laine C, Goldman L, Soukup JR, Hayes JG.** The impact of a regulation restricting medical house staff working hours on the quality of patient care. *JAMA*. 1993;269(3):374-8.
5. **Drazen JM.** Awake and informed. *N Engl J Med*. 2004;351(18):884.
6. **Mukherjee S.** A precarious exchange. *N Engl J Med*. 2004;351(18):1822-4.
7. **Ofri D.** Residency regulations—resisting our reflexes. *N Engl J Med*. 2004;351(18):1824-6.
8. **Silber JH, Williams SV, Krakauer H, Schwartz JS.** Hospital and patient characteristics associated with death after surgery: A study of adverse occurrence and failure-to-rescue. *Med Care*. 1992;30(7):615-29.
9. **Silber JH, Rosenbaum PR, Ross RN.** Comparing the contributions of groups of predictors: Which outcomes vary with hospital rather than patient characteristics? *J Am Stat Assoc*. 1995;90(429):7-18.
10. **Silber JH, Rosenbaum PR, Schwartz JS, Ross RN, Williams SV.** Evaluation of the complication rate as a measure of quality of care in coronary artery bypass graft surgery. *JAMA*. 1995;274(4):317-23.
11. **Silber JH, Romano PS, Rosen AK, Wang Y, Even-Shoshan O, Volpp KG.** Failure-to-rescue: Comparing definitions to measure quality of care. *Med Care*. 2007; 45(10):918-25.
12. **Elixhauser A, Steiner C, Harris DR, Coffey RM.** Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1):8-27.
13. **Glance LG, Dick AW, Osler TM, Mukamel DB.** Does date stamping ICD-9-CM codes increase the value of clinical information in administrative data? *Health Serv Res*. 2006;41(1):231-51.
14. **Quan H, Sundararajan V, Halfon P, et al.** Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005;43(11):1130-9.
15. **Southern DA, Quan H, Ghali WA.** Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. *Med Care*. 2004;42(4):355-60.
16. **Stukenborg GJ, Wagner DP, Connors AF Jr.** Comparison of the performance of two comorbidity measures, with and without information from prior hospitalizations. *Med Care*. 2001;39(7):727-39.
17. **Deyo R, Cherkin D, Ciol M.** Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol*. 1992;45:613-619.
18. **Silber J, Rosenbaum P, Schwartz J, Ross R.** Comparing the contributions of groups of predictors: which outcomes vary with hospital rather than patient characteristics? *J Am Stat Assoc*. 1995;90(429):7-18.
19. **Haberman SJ.** Generalized residuals for log-linear models. *Proceedings of the 9th International Biometric Conference*. 1st ed. Boston: The Biometric Society; 1976:104-23.
20. **Haberman SJ.** The analysis of frequency data. Chicago: The University of Chicago Press; 1974.
21. **Ayanian JZ, Weissman JS.** Teaching hospitals and quality of care: a review of the literature. *Milbank Q*. 2002;80(3):569-93.
22. **Keeler EB, Rubenstein LV, Kahn KL, et al.** Hospital characteristics and quality of care. *JAMA*. 1992;268(13):1709-14.
23. **Allison JJ, Kiefe CI, Weissman NW, et al.** Relationship of hospital teaching status with quality of care and mortality for Medicare patients with acute MI. *JAMA*. 2000;284(10):1256-62.
24. **Taylor DH, Whellan DJ, Sloan FA.** Effects of admission to a teaching hospital on the cost and quality of care for medicare beneficiaries. *N Engl J Med*. 1999;340(4):293-9.
25. **Cox DR.** Note on grouping. *J Am Stat Assoc*. 1957;52(280):543-7.
26. **Cochran WG.** The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*. 1968;24(2):295-313.
27. **Campbell DT, Stanley JC.** Experimental and quasi-experimental designs for research. Dallas: Houghton Mifflin Company; 1963.
28. **Shadish WR, Cook TD, Campbell DT.** Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton-Mifflin; 2002.
29. **Rosenbaum PR.** Stability in the absence of treatment. *J Am Stat Assoc*. 2001;96:210-9.
30. **Lawthers AG, McCarthy EP, Davis RB, Peterson LE, Palmer RH, Iezzoni LI.** Identification of in-hospital complications from claims data. Is it valid? *Med Care*. 2000;38:785-95.
31. **McCarthy EP, Iezzoni LI, Davis RB, et al.** Does clinical evidence support ICD-9-CM diagnosis coding of complications? *Med Care*. 2000;38:868-76.
32. **Weingart SN, Iezzoni LI, Davis RB, et al.** Use of administrative data to find substandard care. Validation of the complications screening program. *Med Care*. 2000;38:796-806.
33. California Office of Statewide Health Planning and Development. (OSHPD). Second report of the California Hospitals Outcomes Project. Acute myocardial infarction. 1996 May.
34. Institute of Medicine. Optimizing graduate medical trainee (resident) hours and work schedules to improve patient safety. 2007.
35. **Myers JS, Bellini LM, Morris JB, et al.** Internal medicine and general surgery residents' attitudes about the ACGME duty hours regulations: A multicenter study. *Acad Med*. 2006;81(12):1052-8.
36. **Jagsi R, Shapiro J, Weissman JS, Dorer DJ, Weinstein DF.** The educational impact of ACGME limits on resident and fellow duty hours: A pre-post survey study. *Acad Med*. 2006;81:1059-68.
37. **Rosen AK, Loveland SA, Romano PS, et al.** Effects of resident duty hour reform on patient safety among hospitalized VA and medicare patients. *Med Care*. 2009. in press.
38. **Silber J, Rosenbaum PR, Rosen AK, et al.** Prolonged hospital stays and the resident duty hour rules of 2003. *Med Care*. 2009. in press.
39. **Landrigan CP, Rothschild JM, Cronin JW, et al.** Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med*. 2004;351(18):1838-48.
40. **Lockley SW, Cronin JW, Evans EE, et al.** Effect of reducing interns' weekly work hours on sleep and attentional failures. *N Engl J Med*. 2004;351(18):1829-37.