Teaching Hospital Five-Year Mortality Trends in the Wake of Duty Hour Reforms

Kevin G. Volpp, MD, PhD^{1,2,3,4}, Dylan S. Small, PhD⁵, Patrick S. Romano, MD, MPH⁶, Kamal M. F. Itani, MD⁷, Amy K. Rosen, PhD^{8,9}, Orit Even-Shoshan, MS^{10,12}, Yanli Wang, MS¹⁰, Lisa Bellini, MD², Michael J. Halenar, BA^{1,2}, Sophia Korovaichuk, BA¹⁰, Jingsan Zhu, MBA^{2,4}, and Jeffrey H. Silber, MD, PhD^{3,10,11,12}

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

BACKGROUND: The Accreditation Council for Graduate Medical Education (ACGME) implemented duty hour regulations for residents in 2003 and again in 2011. While previous studies showed no systematic impacts in the first 2 years post-reform, the impact on mortality in subsequent years has not been examined.

OBJECTIVE: To determine whether duty hour regulations were associated with changes in mortality among Medicare patients in hospitals of different teaching intensity after the first 2 years post-reform.

DESIGN: Observational study using interrupted time series analysis with data from July 1, 2000 to June 30, 2008. Logistic regression was used to examine the change in mortality for patients in more versus less teaching-intensive hospitals before (2000–2003) and after (2003–2008) duty hour reform, adjusting for patient comorbidities, time trends, and hospital site.

PATIENTS: Medicare patients ($n=13,678,956$) admitted to short-term acute care non-federal hospitals with principal diagnoses of acute myocardial infarction (AMI), gastrointestinal bleeding, or congestive heart failure (CHF); or a diagnosis-related group (DRG) classification of general, orthopedic, or vascular surgery.

MAIN MEASURE: All-location mortality within 30 days of hospital admission.

KEY RESULTS: In medical and surgical patients, there were no consistent changes in the odds of mortality at more vs. less teaching intensive hospitals in post-reform years 1–3. However, there were significant relative improvements in mortality for medical patients in the fourth and fifth years post-reform: Post4 (OR 0.88, 95 % CI [0.93–0.94]); Post5 (OR 0.87, [0.82–0.92]) and for surgical patients in the fifth year post-reform: Post5 (OR 0.91, [0.85–0.96]).

CONCLUSIONS: Duty hour reform was associated with no significant change in mortality in the early years after implementation, and with a trend toward improved mortality among medical patients in the fourth and fifth years. It is unclear whether improvements in outcomes long after implementation can be attributed to the reform, but concerns about worsening outcomes seem unfounded.

KEY WORDS: patient outcomes; mortality; duty hour reform; ACGME; administrative data.

J Gen Intern Med 28(8):1048–55

DOI: 10.1007/s11606-013-2401-9

© Society of General Internal Medicine 2013

INTRODUCTION

Widespread concern about effects of resident fatigue, including deaths from medical errors,¹ prompted the Accreditation Council for Graduate Medical Education (ACGME) to implement regulations limiting number of work hours, effective July 1, 2003, for all ACGME-accredited residency programs.^{2,3}

Previous research found that the duty hour regulations had little impact on adult mortality^{4–6} 2 years after implementation, even among high-risk patients,⁷ and minimal or inconsistent impact on indicators of patient safety,⁸ readmission rates,⁹ failure-to-rescue rates (e.g. death given occurrence of a complication), or the probability of prolonged lengths of stay.¹⁰ The ACGME proposed new duty hour standards in July 2011¹¹ that maintain the 80-hour cap on weekly duty hours, but

limit intern shifts to 16 hours. Part of the rationale for this change was that the earlier standards were not adequate in improving patient safety and reducing mortality, and tighter regulation would lead to better outcomes.

The long-term impact of the 2003 duty hour regulations on patient outcomes has not been analyzed. This knowledge gap is important, because some programs did not comply with the regulations until it became clear that they would be strictly enforced.^{12,13} Additionally, if the regulations had beneficial or detrimental effects on resident education, by increasing self-study time or reducing procedural experience,^{14,15} then changes in patient outcomes may not have been realized until the residents who began training under these regulations became supervisors (i.e., 3rd year or later). Finally, if negative effects of the regulations were observed in the first 2 years following implementation, hospitals may have instituted mitigating policies starting in the third year post-reform. In this study, we examine the impact of the 2003 duty hour regulations on mortality for medical and surgical patients enrolled in Medicare and treated at non-federal hospitals over the subsequent 5 years.

METHODS

Approval for this study was obtained from the Institutional Review Boards of The University of Pennsylvania and The Children's Hospital of Philadelphia.

Outcome Measures

The outcome measure was death within 30 days of hospital admission for all patients admitted for acute myocardial infarction (AMI), gastrointestinal bleeding, congestive heart failure (CHF), general surgery, orthopedic surgery, or vascular surgery.¹⁶ There is evidence that mortality varies substantially across institutions for these conditions, and that high mortality may be associated with deficiencies in the quality of care.^{17,18}

Study Sample

Medicare patients were selected using an approach described in previous studies,⁴ including all Medicare patients admitted to short-term general non-federal acute care hospitals from July 1, 2000 to June 30, 2008 with a principal diagnosis of AMI, gastrointestinal bleeding, CHF or with a diagnosis-related group (DRG) classification of general, orthopedic, or vascular surgery. Stroke, included in earlier analyses of duty hour effects, was excluded because stroke mortality increased before the 2003 regulations in teaching hospitals compared to non-teaching hospitals.⁴ The initial sample included 18,331,795 patients from 5,045

acute care hospitals that contributed data for all 8 years within the 50 states or Washington DC.

We excluded patients from hospitals that had fewer than 350 Medicare admissions in any year ($n=45,947$ patients, 933 hospitals) in order to eliminate hospitals that may not have been acute care facilities and those too small to yield stable estimates in a fixed effects analysis. We also excluded patients from hospitals that did not have Medicare Cost Report data ($n=71,947$ patients, 115 hospitals); or that were missing more than 3 months of data in the pre-reform period or 5 months of data in the post-reform periods ($n=543,223$ patients, 938 hospitals).

We also excluded patients younger than 66 years of age to allow 180-day look back for risk adjustment ($n=2,455,361$), or over 90 years of age ($n=1,075,413$), because the changes in the proportion of such patients treated aggressively may not be reflected in administrative data; patients whose hospitalizations spanned July 1, 2003 ($n=28,557$); patients who were enrolled in an HMO at any time during the study period ($n=326,412$); patients who had a reported date of death before an admission date ($n=752$); or patients who were transferred in from another hospital for the qualifying admission ($n=23,509$) to avoid double counting admissions within 30 days. Among AMI patients, those discharged alive in fewer than 2 days ($n=81,718$) were excluded because such cases may not represent actual AMIs.¹⁹⁻²¹ All admissions for patients who met the above conditions were included, leaving a sample of 13,678,956 patients from 3,059 hospitals.

Risk Adjustment and Data

The risk-adjustment approach was identical to our previous reports,^{4,5,7-10} a modification of methods of Elixhauser et al.²²

Data on patient characteristics were obtained from the Medicare Provider Analysis and Treatment File (MEDPAR), which includes information on principal and secondary diagnoses, age, sex, comorbidities, and discharge status, including dates of death. Data on health maintenance organization (HMO) enrollment were obtained from the Center for Medicare and Medicaid Services (CMS) denominator files. The number of residents at each hospital was obtained from CMS Medicare Cost Reports.²³

The primary measure of teaching intensity was the resident-to-bed ratio, calculated at a defined point in time as (number of interns + residents)/mean number of staffed beds. The resident-to-bed ratio has been used to differentiate major teaching, minor teaching, and non-teaching hospitals in previous studies,²⁴⁻²⁶ and we used a similar approach in previous work.^{4,5,8-10}

Statistical Analysis

We used a multiple time series research design,²⁷ also known as "difference-in-differences," to examine whether

the change in duty hour rules was associated with a change in the underlying trend in patient outcomes in teaching hospitals, an approach that reduces potential biases from unmeasured variables.^{28,29} The multiple time series research design compares each hospital to itself, contrasting the changes in hospitals with more residents to the changes in hospitals with fewer or no residents, and adjusting for observed differences in patient risk factors. The design also adjusts for changes in outcomes over time (trends) common to all hospitals and minimizes bias from three possible sources. First, a difference between hospitals that is stable over time cannot be mistaken for an effect of the reform because hospital fixed effects are used to compare each hospital to itself, before and after reform. Second, universal changes affecting all hospitals similarly, such as technological improvements, cannot be mistaken for an effect of the reform because the logit model includes year indicators. Third, if the mix of patients is changing differently among hospitals, and if these changes are accurately reflected in measured risk factors, this cannot be mistaken for an effect of the reform because the logit model adjusts for these measured risk factors.

Notwithstanding these advantages, the difference-in-differences method has limitations. Any diverging trend in mortality over time for more vs. less teaching-intensive hospitals already in progress or coincident with the initiation of the reform could be mistaken for an effect of the reform, although we adjusted for any observed differences in pre-reform trends. Less teaching-intensive hospitals, including all non-teaching hospitals, served as the primary control group 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.’ They are also geographically diverse with large patient populations, and similar patient discharge data are available. Data from July 1, 2000 to June 30, 2003 were designated the pre-reform period, with data from July 1, 2003 to June 30, 2008 designated the post-reform period.

Whereas our previous analyses utilized one observation per patient over a 5-year period,^{4,5,7–10} we now utilize all admissions since multiple admissions per patient are more common during the 8-year time period of this study.³⁰ To estimate confidence intervals and p values that account for the correlation between patient observations, we utilized the bootstrap procedure.^{28,29,31}

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. The effect of the change in duty hour rules was estimated using the coefficients of resident-to-bed ratio interacted with dummy variables indicating post-reform years 1, 2, 3, 4, and 5 (e.g. post-reform year 1 was from

July 1, 2003 to June 30, 2004). These coefficients, presented as odds ratios, measure the degree to which mortality changed in more vs. less teaching-intensive hospitals 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 both individually and together as “combined medical” and “combined surgical” groups.

We used a similar modeling strategy as in our previous work.^{4,5,7–10} Briefly, risk-adjusted mortality was assumed to have a common time trend until implementation of the duty hour rules, after which the trend at teaching hospitals 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 duty hour reform, a “test of controls” was performed. Parameters were added to the model for interactions between the resident-to-bed ratio and indicators for pre-reform year 2 and pre-reform year 1. A Wald test was used to determine whether these interactions significantly improved model fit, suggesting that teaching and non-teaching hospitals had differing trends in mortality in the 3 years pre-reform independent of the reform. When such differing trends were found, post hoc analyses were conducted in which post-reform results were compared to pre-reform year 1 as a baseline instead of the entire 3-year pre-reform period.

RESULTS

The number of admissions for each of the conditions was fairly constant over time, differing by less than 10 % per year for any condition (Table 1). About 68 % of the hospitals were non-teaching hospitals, and about 9.8 % of hospitals (treating 14.6 % of patients) were major or very major teaching hospitals (Table 2).

Between pre-reform year 3 (academic year [AY] 2000–2001) and pre-reform year 1 (AY 2002–2003), unadjusted mortality rates for the combined medical group changed at similar rates among hospitals of different teaching intensity, indicating that less teaching intensive hospitals can reasonably serve as a control group in analyzing mortality trends at more teaching intensive hospitals (Fig. 1a). Unadjusted mortality for the combined medical group changed at similar rates for patients in hospitals of differing teaching intensity between pre-reform year 3 and post-reform year 3 (AY 2005–06), with a possible trend towards improved mortality in the very major teaching hospitals in post-reform years 4 and 5 (AY 2006–2008) (Fig. 1a). Among surgical patients, mortality appeared to change at similar rates in hospitals of different teaching intensity from pre-reform year 3 (AY 2000–2001) to post-reform year 5 (AY

Table 1. Unadjusted Mortality Rates by Condition and Year Relative to the Implementation Date of ACGME Duty Hour Regulations in July 2003

		Pre 3 (2000–1)	Pre 2 (2001–2)	Pre 1 (2002–3)	Post 1 (2003–4)	Post 2 (2004–5)	Post 3 (2005–6)	Post 4 (2006–7)	Post 5 (2007–8)
Medical conditions									
AMI	Number of cases	255,233	252,681	248,769	243,598	232,161	209,587	197,740	190,069
	Unadjusted mortality rate (%)	16.7	16.4	15.7	15.6	15.1	14.2	14.0	13.9
GI bleed	Number of cases	212,268	218,030	220,477	229,212	226,479	214,507	206,082	194,918
	Unadjusted mortality rate (%)	6.8	6.7	6.5	6.3	6.2	5.9	6.0	6.1
CHF	Number of cases	531,184	530,174	544,289	554,723	554,603	519,508	486,325	441,464
	Unadjusted mortality rate (%)	10.0	9.9	9.5	9.5	9.5	9.0	9.2	9.3
Surgical categories									
General surgery	Number of cases	432,402	437,357	434,136	438,720	431,404	418,269	398,388	390,619
	Unadjusted mortality rate (%)	5.3	5.3	5.3	5.3	5.2	5.0	4.9	4.9
Orthopedic surgery	Number of cases	645,347	683,922	713,957	768,859	804,021	798,165	793,999	780,904
	Unadjusted mortality rate (%)	2.4	2.3	2.2	2.1	2.0	1.9	1.9	1.9
Vascular surgery	Number of cases	98,791	98,659	96,268	94,115	91,953	87,085	81,444	74,649
	Unadjusted mortality rate (%)	12.3	12.4	12.1	11.9	11.6	11.1	10.9	10.9

AMI acute myocardial infarction, GI gastrointestinal, CHF congestive heart failure

2007–8) (Fig. 1b). For both medical and surgical groups, the test of controls was rejected and all three pre-reform years were used as the reference group in the analyses.

In adjusted analyses of the combined medical group (Table 3), slight relative improvements in mortality in post-reform years 1 (OR 0.95, 95 % CI [0.91, 0.99]) and 3 (OR 0.94, 95 % CI [0.89, 0.99]) became more pronounced in years 4 (OR 0.88, 95 % CI [0.83, 0.94]) and 5 (OR 0.87, 95 % CI [0.82, 0.92]). The condition specific analyses for AMI, GI bleed, and CHF follow the same general pattern, with the relative improvement in mortality in more teaching intensive hospitals becoming apparent for all three conditions by post-reform year 5. In adjusted analyses of the combined surgical patients, we observe no relative changes in mortality in accordance with teaching intensity in post-reform years 1 and 4, but significant improvement in year 5 (OR 0.91, 95 % CI [0.85, 0.96]).

Excluding patients admitted to hospitals in New York State (a state that had instituted resident hour reform prior to our reference years) or patients admitted from nursing homes (whose care may be less aggressive) did not change the results. We also examined whether changes in the coding of comorbidities could explain any of these effects. There was only a 0.8 % relative decrease in the mean number of comorbidities in more teaching-intensive hospitals relative to non-teaching hospitals from pre-reform year 1 to post-reform year 5, and sensitivity analyses without adjustment for comorbidities produced similar results. As in previous work,⁷ we did not find any significant differences in the degree to which mortality changed in more vs. less

teaching intensive hospitals in any of the five post reform years for high vs. low-risk patients for either the combined medical group or the combined surgical group (results in Appendix).

To illustrate the magnitude of the changes in mortality, we estimated the adjusted risk of mortality in each academic year for a hypothetical patient with mean values of all regression covariates at two hypothetical hospitals: a non-teaching hospital with a resident-to-bed ratio of 0, and a highly teaching intensive hospital with a resident-to-bed ratio of 1. In the combined medical group, mortality in a hospital with a resident-to-bed ratio of 1 decreased from 7.56 % in pre-reform year 1 to 5.48 % in post-reform year 5, while in a hospital with a resident-to-bed ratio of 0 mortality decreased from 9.09 % to 7.47 %, a relative decrease of 0.46 percentage points or 5.06 %. For a patient representing the combined surgical group, mortality in a hospital with a resident-to-bed ratio of 1 decreased from 1.50 % in pre-reform year 1 to 1.15 % in post-reform year 5, while in a hospital with a resident-to-bed ratio of 0, mortality decreased from 2.06 % to 1.76 %, a relative decrease of 0.04 percentage points or 1.94 %.

COMMENTS

Implementation of the 2003 duty hour rules was arguably one of the largest efforts undertaken to reduce errors in teaching hospitals. The results of this study confirm that there were no systematic effects on mortality in the first

Table 2. Hospitals and Admissions by Teaching Status

	Nonteaching (0)	Very minor teaching (> 0–0.049)	Minor teaching (0.050–0.249)	Major teaching (0.250–0.599)	Very major teaching (0.600–1.090)
Number (%) of hospitals*	2070 (67.7 %)	296 (9.7 %)	394 (12.9 %)	193 (6.3 %)	106 (3.5 %)
Number (%) of admissions	9,819,358 (49.9 %)	2,818,337 (14.3 %)	4,153,529 (21.1 %)	1,904,338 (9.7 %)	967,884 (4.9 %)

*Total number of hospitals included is 3,059

Figure 1. a Unadjusted mortality for combined medical patients by teaching-intensive hospital status. **b** Unadjusted mortality for combined surgical patients by teaching-intensive hospital status.

3 years post-reform, but suggest a relative improvement in mortality at high teaching intensity hospitals by the 5th year post-reform among both medical and surgical patients. The absence of relative changes in mortality in the first 3 years post-reform suggests that the duty hour regulations did not have any direct effect on mortality. The outcomes in high-risk patients were similarly unaffected in the first 3 years, further suggesting that the regulations had no direct impact on mortality. Improvements in mortality in post-reform years 4 and 5 suggest that the reform did not harm patients, although we are hesitant to attribute these delayed effects to

the reform since it cannot be determined whether the improvements would have been smaller or greater had no reform been implemented.

Other major concurrent policy changes of differential impact on teaching and non-teaching hospitals could play a role in the observed outcomes. For example, Section 501(b) of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003 authorized CMS to penalize hospitals that did not report designated quality measures by reducing their annual market basket update. The Deficit Reduction Act of 2005 increased that reduction

Table 3. Odds of Mortality in More Versus Less Teaching Intensive Hospitals Post Duty Hour Reform of 2003

Patient categories	No. of cases	RB x post 1	RB x post 2	RB x post 3	RB x post 4	RB x post 5
Medical conditions						
AMI	280,337/ 1,829,838	0.98 (0.90, 1.08) P = 0.71	1.01 (0.93, 1.10) P = 0.74	1.04 (0.95, 1.14) P = 0.37	0.96 (0.88, 1.05) P = 0.38	0.91 (0.83, 0.999) P = 0.047
GI bleed	108,690/ 1,721,973	1.06 (0.95, 1.19) P = 0.31	0.87 (0.77, 0.98) P = 0.02	1.00 (0.98, 1.14) P = 0.97	0.87 (0.77, 0.99) P = 0.03	0.86 (0.76, 0.97) P = 0.02
CHF	395,689/ 4,162,270	0.92 (0.86, 0.99) P = 0.03	1.04 (0.97, 1.10) P = 0.28	0.92 (0.86, 0.98) P = 0.01	0.89 (0.83, 0.95) P = 0.0007	0.85 (0.79, 0.92) P < 0.0001
Combined medical	784,716/ 7,714,081	0.95 (0.91, 0.99) P = 0.01	0.98 (0.93, 1.02) P = 0.25	0.94 (0.89, 0.99) P = 0.01	0.88 (0.83, 0.94) P = 0.0002	0.87 (0.82, 0.92) P < 0.0001
Surgical categories						
General surgery	174,177/ 3,381,295	1.03 (0.97, 1.10) P = 0.30	0.98 (0.90, 1.07) P = 0.72	0.96 (0.88, 1.05) P = 0.36	0.93 (0.85, 1.01) P = 0.10	0.84 (0.77, 0.92) P = 0.0001
Orthopedic surgery	123,638/ 5,989,174	1.05 (0.91, 1.20) P = 0.53	1.02 (0.91, 1.13) P = 0.78	1.05 (0.93, 1.18) P = 0.47	0.95 (0.85, 1.06) P = 0.34	0.94 (0.83, 1.06) P = 0.30
Vascular surgery	84,448/ 722,964	0.99 (0.91, 1.08) P = 0.84	1.00 (0.89, 1.12) P = 0.99	1.02 (0.91, 1.14) P = 0.71	1.05 (0.94, 1.17) P = 0.39	1.05 (0.94, 1.17) P = 0.38
Combined surgical	382,263/ 10,093,433	1.02 (0.97, 1.08) P = 0.48	1.00 (0.94, 1.06) P = 0.89	0.98 (0.93, 1.04) P = 0.59	0.95 (0.89, 1.00) P = 0.07	0.91 (0.85, 0.96) P = 0.002

The interaction terms (e.g. resident-to-bed ratio $\times$ post-reform year 1) measure whether there is any relative change in the odds of mortality in more vs. less teaching intensive hospitals from pre-reform to each respective year following the ACGME duty hour reform of 2003. For example, Post 1 refers to academic year 2003–4, post 2 academic year 2004–5. The coefficients shown are bootstrapped odds ratios and 95 % confidence intervals for risk-adjusted mortality at high-teaching intensity hospitals (resident-to-bed ratio of 1.0), relative to low teaching-intensity hospitals (resident-to-bed ratio of 0), by year after implementation of 2003 ACGME duty hour regulations. An odds ratio of greater than 1.0 indicates a relative increase in mortality in more vs. less teaching intensive hospitals from the pre-reform period to that particular post-reform year. Odds ratios significantly different than 1.0 are highlighted in yellow. The time period from 2000 to 2003 (Pre years 1–3) was used as the baseline period for all of these analyses.

from 0.4 % to 2.0 %, leading to a rapid increase in participation.³¹ During the same period, 225 hospitals joined the Premier Hospital Quality Incentive Demonstration, which was established to reward hospitals demonstrating high quality in several of the clinical domains studied here: AMI, CHF, pneumonia, and surgical care.³² The MMA improved the financial attractiveness of Critical Access Hospital designation, such that 1,327 rural non-teaching hospitals have now pulled out of the Prospective Payment System and its attendant public reporting requirements.³³ With respect to surgical care, concurrent interventions include the roll-out of the National Surgical Quality Improvement program (NSQIP) from the VA to the private sector and the implementation of The Joint Commission's Surgical Care Improvement Project (SCIP). Although nationwide initiatives, it is possible that they were differentially adopted by hospitals at different levels of teaching intensity. These concurrent changes make it impossible to determine the incremental impact of any

single intervention, but we can clearly say that mortality did not worsen following implementation of the 2003 duty hour rules.

The absence of any clinically important relative change in mortality in the first 3 years post-reform in more teaching intensive hospitals may reflect many factors. It is possible that mortality increased following the rule change, and that any negative effects of the reform were offset by technological improvements or more efficient management of handoffs. Such factors cannot be disentangled given the nature of observational data; we can only study the net change. Lack of compliance with the duty hour rules has also been described, although such surveys are limited to the self-reports of a small and potentially non-representative sample of about 7 % of all interns,³⁴ or data collected by the ACGME.³⁵ Strong financial incentives to avoid loss of accreditation deter widespread resistance to the regulations. Unless appropriate institutional resources were put in place to mitigate the impact of reducing resident work hours, the

effect of the reform at the local level would likely be to increase work intensity, which could offset potential benefits of decreased fatigue.

There are several limitations to our study. We included in part an LOS-based definition of AMI historically used to rule out misclassifications of AMI.^{19–21} Further, we focused on one outcome, mortality. While the duty hour rules were an attempt to reduce deaths from medical errors, measurement of other outcomes such as complications may elucidate the relative effects of decreased continuity of care compared to decreased resident fatigue. Since we do not have information on hours worked by each resident, our study should be considered an effectiveness, not an efficacy, study because we measured the outcomes associated with the duty hour rules as implemented. Even with the size of the Medicare population, some of the confidence intervals are still broad, and we cannot rule out small but clinically meaningful effects. Any observational study is susceptible to unmeasured confounding, which is a particular challenge given the duration of this study. We used administrative data, so risk adjustment is more limited than with clinical data; however, by comparing outcomes over time within each hospital, in more versus less teaching-intensive hospitals, potential bias from unmeasured confounders is markedly diminished. To be a confounder in our difference-in-differences analysis, an unmeasured risk factor must have changed at different rates in more versus less teaching-intensive hospitals. However, survival bias associated with admissions resulting in death being differentially more likely post-reform than pre-reform in teaching hospitals vs. non-teaching hospitals could still affect this study's results.⁴

In summary, implementation of duty hour regulations was associated with a significant improvement in risk-adjusted mortality for medical patients in more teaching intensive hospitals in post-reform years 4 and 5 and among surgical patients in post-reform year 5. These results do not address whether the design of duty hour rules is optimal.³⁶ However, they do suggest that the 2003 regulations had little initial effect on risk-adjusted mortality, but that quality, as measured by 30-day mortality, did begin to improve at greater rates in more teaching intensive hospitals several years thereafter. Further empirical work should examine whether similar patterns are observed in conjunction with the duty hour reform of 2011.

Acknowledgements: This work was funded by grant R01 HL094593-01 from the NHLBI.

Conflict of Interest: 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. There are no known financial conflicts of interest among any of the authors, including but not limited to employment/affiliation, grants or funding, honoraria, paid consultancies, expert testimony, stock ownership or options, and patents filed, received or pending. Both Dr. Volpp and Dr. Bellini served as unpaid members of the Committee on Innovations for the ACGME from 2005 to 2009.

Corresponding Author: Kevin G. Volpp, MD, PhD; Center for Health Equity Research and Promotion, Veteran's Administration Hospital, 1120 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, USA (e-mail: Volpp70@wharton.upenn.edu).

REFERENCES

1. Kohn LT, Corrigan JM, Donaldson MS. To Err is Human: Building a Safer Health System. Washington: National Academy Press; 2000.
2. Accreditation Council for Graduate Medical Education. Resident duty hours language: Final requirements. 2003; <http://www.acgme.org>. Accessed February 18, 2013.
3. Accreditation Council for Graduate Medical Education. Report of the work group on resident duty hours and the learning environment. June 11, 2002. In: The ACGME's approach to limit resident duty hours 12 months after implementation: a summary of achievements, 2004. 2006; http://www.acgme.org/acgmeweb/Portals/0/PFAssets/PublicationsPapers/dh_dutyhourssummary2003-04.pdf. Accessed February 18, 2013.
4. Volpp KG, Rosen AK, Rosenbaum PR, et al. Mortality among hospitalized Medicare beneficiaries in the first 2 years following ACGME resident duty hour reform. *JAMA*. 2007;298(9):975–83.
5. Volpp KG, Rosen AK, Rosenbaum PR, et al. Mortality among patients in VA hospitals in the first 2 years following ACGME resident duty hour reform. *JAMA*. 2007;298(9):984–92.
6. Shetty KD, Bhattacharya J. Changes in hospital mortality associated with residency work-hour regulations. *Ann Intern Med*. 2007;147(2):73–80.
7. Volpp KG, Rosen AK, Rosenbaum PR, et al. Did duty hour reform lead to better outcomes among the highest risk patients? *J Gen Intern Med*. 2009;24(10):1149–55.
8. Rosen AK, Loveland SA, Romano PS, et al. Effects of resident duty hour reform on surgical and procedural patient safety indicators among hospitalized Veterans Health Administration and Medicare patients. *Med Care*. 2009;47(7):723–31.
9. Press MJ, Silber JH, Rosen AK, et al. The impact of resident duty hour reform on hospital readmission rates among Medicare beneficiaries. *J Gen Intern Med*. 2011;26(4):405–11.
10. Silber JH, Rosenbaum PR, Rosen AK, et al. Prolonged hospital stay and the resident duty hour rules of 2003. *Med Care*. 2009;47(12):1191–200.
11. Nasca TJ, Day SH, Amis ES, Jr. The new recommendations on duty hours from the ACGME Task Force. *N Engl J Med*. 2010.
12. Croasdale M. Johns Hopkins Penalized for Resident Hour Violations. 2003; <http://www.ama-assn.org/amednews/2003/09/15/prsc0915.htm>. Accessed February 18, 2013.
13. Killelea BK, Chao L, Scarpinato V, Wallack MK. The 80-hour workweek. *Surg Clin N Am*. 2004;84(6):1557–72, x.
14. Fletcher KE, Reed DA, Arora VM. Patient safety, resident education and resident well-being following implementation of the 2003 ACGME duty hour rules. *J Gen Intern Med*. 2011;26(8):907–19.
15. Moonesinghe SR, Lowery J, Shahi N, Millen A, Beard JD. Impact of reduction in working hours for doctors in training on postgraduate medical education and patients' outcomes: systematic review. *BMJ (Clin Res Ed)*. 2011;342:d1580.
16. Iezzoni LI. Risk Adjustment for Measuring Health-care Outcomes. 3rd ed. Chicago: Health Administration Press; 2003.
17. Rockall TA, Logan RF, Devlin HB, Northfield TC. Variation in outcome after acute upper gastrointestinal haemorrhage. The National Audit of Acute Upper Gastrointestinal Haemorrhage. *Lancet*. 1995;346(8971):346–50.
18. Agency for Healthcare Research and Quality. Inpatient Quality Indicators Download. AHRQ Quality Indicators. 2007; http://www.qualityindicators.ahrq.gov/downloads/software/sas/v21r4/iqui_guide_rev4.pdf Accessed February 18, 2013.
19. Second Report of the California Hospitals Outcomes Project. Acute Myocardial Infarction 1996; <http://repositories.cdlib.org/chsrpc/coshpd/index.html#1996>. Accessed February 18, 2013.
20. Iezzoni LI, Burnside S, Sickles L, Moskowitz MA, Sawitz E. Coding of acute myocardial infarction: clinical policy and implications. *Ann Intern Med*. 1988;109:745–51.
21. Kiyota Y, Schneeweiss S, Glynn RJ, Cannuscio CC, Avorn J, Solomon DH. Accuracy of Medicare claims-based diagnosis of acute myocardial infarction: estimating positive predictive value on the basis of review of hospital records. *Am Hear J*. 2004;148(1):99–104.

22. **Elixhauser A, Steiner C, Harris DR, Coffey RM.** Comorbidity measures for use with administrative data. *Med Care.* 1998;36(1):8–27.
23. **Lawthers AG, McCarthy EP, Davis RB, Petersen LE, Palmer RH, Iezzoni LI.** Identification of in-hospital complications from claims data: is it valid? *Med Care.* 2000;38(8):785–95.
24. **Keeler EB, Rubenstein LV, Kahn KL, et al.** Hospital characteristics and quality of care. *JAMA.* 1992;268(13):1709–14.
25. **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.
26. **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.
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. **Field C, Welsh A.** Bootstrapping clustered data. *J R Stat Soc Ser B Methodol.* 2007;69:369–90.
31. CMS. Fiscal year 2009 quality measure reporting for 2010 payment update 2009; <https://www.cms.gov/HospitalQualityInits/downloads/HospitalRHQDAPU200808.pdf>. Accessed February 18, 2013.
32. CMS. Premier Hospital Quality Incentive Demonstration. 2011; https://www.cms.gov/HospitalQualityInits/35_HospitalPremier.asp. Accessed February 18, 2013.
33. RAC. Critical access hospital (CAH) frequently asked questions. Health and human services information for rural America 2011; <http://www.raonline.org/topics/hospitals/cahfaq.php>. Accessed February 18, 2013.
34. **Landrigan CP, Barger LK, Cade BE, Ayas NT, Czeisler CA.** Interns' compliance with accreditation council for graduate medical education work-hour limits. *JAMA.* 2006;296(9):1063–70.
35. Accreditation Council for Graduate Medical Education. The ACGME's approach to limit resident duty hours 12 months after implementation: A summary of achievements, 2004. 2006; http://www.acgme.org/acgmeweb/Portals/0/PFAssets/PublicationsPapers/dh_dutyhourssummary2003-04.pdf. Accessed February 18, 2013.
36. **Pellegrini VD Jr, Peabody T, Dinges DF, Moody J, Fabri PJ.** Symposium resident work-hour guidelines. A sentence or an opportunity for orthopaedic education? *J Bone Joint Surg Am.* 2005;87(11):2576–86.

APPENDIX

Table 4. Changes in Odds of Mortality for Higher Risk Patients Versus Lower Risk Patients in Conjunction with Duty Hour Reform Implementation

	RB X post1 X top 10%/25% of severity	RB X post2 X top 10%/25% of severity	RB X post3 X top 10%/25% of severity	RB X post4 X top 10%/25% of severity	RB X post5 X top 10%/25% of severity
Combined medical					
Highest 10% (vs. bottom 90%)	1.05 (0.96, 1.15) P = 0.32	0.96 (0.87, 1.06) P = 0.40	0.98 (0.90, 1.06) P = 0.60	0.95 (0.87, 1.04) P = 0.21	0.92 (0.84, 1.01) P = 0.07
Highest 25% (vs. bottom 75%)	1.00 (0.92, 1.08) P = 0.97	0.96 (0.90, 1.03) P = 0.29	0.99 (0.92, 1.07) P = 0.83	0.97 (0.90, 1.04) P = 0.35	0.93 (0.85, 1.00) P = 0.06
Combined surgical					
Highest 10% (vs. bottom 90%)	0.88 (0.79, 0.99) P = 0.03	1.02 (0.91, 1.14) P = 0.78	1.00 (0.88, 1.13) P = 0.96	1.02 (0.90, 1.15) P = 0.75	0.96 (0.85, 1.10) P = 0.59
Highest 25% (vs. bottom 75%)	1.00 (0.89, 1.12) P = 0.95	1.02 (0.90, 1.16) P = 0.73	1.06 (0.93, 1.21) P = 0.38	1.02 (0.90, 1.15) P = 0.59	0.96 (0.85, 1.10) P = 0.59
FTR rates					
Highest 10% (vs. bottom 90%)	0.94 (0.82, 1.06) P = 0.31	1.03 (0.92, 1.15) P = 0.62	1.07 (0.95, 1.22) P = 0.26	1.04 (0.89, 1.21) P = 0.63	0.97 (0.86, 1.10) P = 0.67
Highest 25% (vs. bottom 75%)	0.96 (0.87, 1.05) P = 0.36	1.01 (0.91, 1.12) P = 0.90	1.02 (0.91, 1.14) P = 0.80	1.03 (0.93, 1.14) P = 0.61	1.03 (0.91, 1.18) P = 0.63

The interaction terms (e.g. resident-to-bed ratio $\times$ post-reform year 1) measure whether there is any relative change in the odds of mortality in more vs. less teaching intensive hospitals. The coefficients shown are bootstrapped odds ratios and 95 % confidence intervals that examine changes in the odds of mortality for the highest risk patients vs. lower risk patients at high-teaching intensity hospitals (resident-to-bed ratio of 1.0), relative to low teaching-intensity hospitals (resident-to-bed ratio of 0), by year after implementation of 2003 ACGME duty hour regulations. An odds ratio of greater than 1.0 indicates a relative increase in mortality among high risk patients in more vs. less teaching intensive hospitals from the pre-reform period to that particular post-reform year. Odds ratios significantly different than 1.0 are highlighted in yellow. The time period from 2000 to 2003 (Pre years 1–3) was used as the baseline period for all of these analyses.