

ONLINE FIRST

Effects of the 2011 Duty Hour Reforms on Interns and Their Patients

A Prospective Longitudinal Cohort Study

Srijan Sen, MD, PhD; Henry R. Kranzler, MD; Aashish K. Didwania, MD; Ann C. Schwartz, MD; Sudha Amarnath, MD; Joseph C. Kolars, MD; Gregory W. Dalack, MD; Breck Nichols, MD, MPH; Constance Guille, MD

Importance: In 2003, the first phase of duty hour requirements for US residency programs recommended by the Accreditation Council for Graduate Medical Education (ACGME) was implemented. Evidence suggests that this first phase of duty hour requirements resulted in a modest improvement in resident well-being and patient safety. To build on these initial changes, the ACGME recommended a new set of duty hour requirements that took effect in July 2011.

Objective: To determine the effects of the 2011 duty hour reforms on first-year residents (interns) and their patients.

Design: As part of the Intern Health Study, we conducted a longitudinal cohort study comparing interns serving before (2009 and 2010) and interns serving after (2011) the implementation of the new duty hour requirements.

Setting: Fifty-one residency programs at 14 university and community-based GME institutions.

Participants: A total of 2323 medical interns.

Main Outcome Measures: Self-reported duty hours, hours of sleep, depressive symptoms, well-being, and medical errors at 3, 6, 9, and 12 months of the internship year.

Results: Fifty-eight percent of invited interns chose to participate in the study. Reported duty hours decreased from an average of 67.0 hours per week before the new rules to 64.3 hours per week after the new rules were instituted ($P < .001$). Despite the decrease in duty hours, there were no significant changes in hours slept ($6.8 \rightarrow 7.0$; $P = .17$), depressive symptoms ($5.8 \rightarrow 5.7$; $P = .55$) or well-being score ($48.5 \rightarrow 48.4$; $P = .86$) reported by interns. With the new duty hour rules, the percentage of interns who reported concern about making a serious medical error increased from 19.9% to 23.3% ($P = .007$).

Conclusions and Relevance: Although interns report working fewer hours under the new duty hour restrictions, this decrease has not been accompanied by an increase in hours of sleep or an improvement in depressive symptoms or well-being but has been accompanied by an unanticipated increase in self-reported medical errors.

JAMA Intern Med.

Published online March 25, 2013.

doi:10.1001/jamainternmed.2013.351

OVER THE PAST 25 YEARS, there has been a growing concern that the long duty hours and sleep deprivation that have traditionally been common during residency training can lead to adverse consequences for both residents and the patients that they treat.¹ How best to promote rigorous, high-quality medical training while maintaining the health of residents and maximizing the safety of patients is intensely debated.²

See Editor's Note

Using input from the Institute of Medicine (IOM), the Accreditation Council for Graduate Medical Education (ACGME) established a new set of duty hour rec-

ommendations effective July 2011.^{1,3} While, the ACGME recommendations include numerous changes in supervision and in required oversight by programs, the most controversial change was to limit the maximum shift length for first-year residents (commonly referred to as interns) to 16 hours. This recommendation was based on studies showing that reducing the length of intern shifts in the intensive care unit reduced the incidence of serious medical errors⁴ and that long duty hours harm both patients^{5,6} and residents themselves.⁷⁻⁹ Limiting shift length is controversial because other research has shown a substantial increase in patient handoffs with shorter shifts and a higher rate of medical errors with more handoffs.^{10,11} Furthermore, a recent study found that

Author Affiliations are listed at the end of this article.

duty hour restrictions increased the stress felt by residents, despite a reduction in extended shifts.¹²

To assess the impact of the 2011 changes we analyzed data from the Intern Health Study,¹³ a large, multispecialty and multi-institutional longitudinal study of interns. In particular, we compared the experiences of interns serving before and interns serving after the new duty hour standards in terms of duty hours, sleep duration, depressive symptoms, and medical errors.

METHODS

PARTICIPANTS

Graduate medical education (GME) institutions and medical schools who expressed interest in taking part in the Intern Health Study were included in the sample for this study.¹³ Fifty-one residency programs at 10 university-based and 4 community-based GME institutions allowed us to invite their incoming residents to participate in the study. In addition, 4 medical schools agreed to invite their graduating students to participate in the study. In total, 4352 individuals entering internal medicine, general surgery, pediatrics, obstetrics and gynecology, emergency medicine, combined medicine and pediatrics, transitional year, and psychiatry residency programs during the 2009, 2010, and 2011 academic years were invited to participate via e-mail sent 2 months prior to commencing internship.

E-mail invitations for 348 individuals were returned as undeliverable, and we were unable to obtain a valid e-mail address. Fifty-eight percent of the remaining invited individuals (2323 of 4005) agreed to participate in the study. The institutional review boards at all of the participating institutions approved the study. Participants were compensated \$50 for their time and effort. Because there was some variation from year to year in the set of residency programs assessed in the study, we performed a secondary analysis including only participants at residency programs that took part in the study in all 3 years of study.

The set of included residency programs varied slightly across the 2009, 2010, and 2011 cohorts for 2 reasons: (1) the participants invited through their graduating medical school enrolled in a different set of residency programs each year and (2) 9 residency programs took part in only 1 or 2 years of the study. To evaluate whether the different set of programs between cohorts was a confounder, we conducted a secondary analysis of the “common program subsample” including only residency programs that took part in the study during all 3 academic years.

DATA COLLECTION

The procedures used in the Intern Health Study have been detailed previously.¹³ Briefly, all surveys were conducted through a secure online website designed to maintain confidentiality, with participants identified only by numbers.

Preinternship Assessment

Participants completed a baseline survey 2 months prior to commencing internship that assessed general demographic factors (age, sex, marital status), and psychological factors (baseline Patient Health Questionnaire–9 [PHQ-9] depressive symptoms, history of depression and a measure of personality, the NEO-Five Factor Inventory Neuroticism).^{14,15} The PHQ-9 is a 9-item self-report component of the Primary Care Evaluation of Mental Disorders (PRIME-MD) inventory, designed to screen for depressive symptoms. A score of 10.0 or greater on the PHQ-9 has a sensi-

tivity of 93% and a specificity of 88% for the diagnosis of major depressive disorder.¹⁶

Within-Internship Assessments

Interns were contacted via e-mail on the first Tuesday of months 3, 6, 9, and 12 of the internship year and asked to complete a previously published survey addressing work hours, sleep, and medical error.¹³ Survey questions included “How many hours have you worked in the past week?” and “On average, how many hours have you slept per night over the past week?” To assess medical errors, the survey asked the question “Are you concerned that you have made any major medical errors in the last 3 months?”¹⁷ Interns’ depressive symptoms were assessed with the PHQ-9. To minimize the time burden on interns, Subjective Well-Being, assessed through the 14-item Mental Health Continuum–Short Form, was administered only at month 6.¹⁸

STATISTICAL ANALYSIS

Baseline factors were compared between the preimplementation and postimplementation cohorts with a 2-sample *t* test for continuous measures and a χ^2 test for categorical measures. Generalized estimating equations (GEE) were used to assess the effects of the duty hour rule while accounting for the repeated measures within participants during internship. In the GEE analyses, cohort membership was entered as the predictor variable (preimplementation [2009 and 2010] cohort compared with postimplementation [2011] cohort), and quarterly reports of duty hours, sleep hours, depressive symptoms, and medical errors were assessed as outcomes. To evaluate differences in response to duty hour reforms among medical specialties, a cohort membership \times specialty interaction term was included in these analyses. Two-sample *t* tests were used to compare subjective well-being between cohorts. Analyses were performed using SPSS software (version 19.0; SPSS Inc).

RESULTS

REPRESENTATIVENESS OF THE SAMPLE

Compared with all individuals entering internship in 2009, 2010, and 2011 nationally, our sample was younger (27.5 years old vs 28.8 years old) and included a slightly higher percentage of women (50.9% vs 48.6%). There were no significant differences in specialty, institution, or demographic variables between individuals who participated and those who did not ($P > .05$ for all comparisons) (demographic information for full residency programs provided by the American Association of Medical Colleges; e-mail communication; September 17, 2012).

PREINTERNSHIP COMPARISONS

For the 2323 study participants, there were no significant differences in specialty, baseline depressive symptom score, neuroticism, sex, or age between the preimplementation (2009 [n=714] and 2010 [n=772]) and postimplementation (2011 [n=837]) cohorts (**Table 1**; all P values $> .05$ [$P = .23-.95$]). Furthermore, there were no differences in these baseline characteristics between participants who completed 1, 2, 3, or 4 quarterly surveys (all P values $> .05$ [$P = .18-.93$]). On average, 65% of participants responded to each quarterly survey. In total,

Table 1. Sample Demographic Characteristics^a

Variable ^b	2009	2010	2011
Sex			
Male	345 (48.3)	383 (49.6)	412 (49.2)
Female	369 (51.7)	389 (50.4)	425 (50.8)
Specialty			
Internal medicine	404 (56.6)	412 (53.4)	451 (53.9)
General surgery	84 (11.8)	67 (8.7)	86 (10.2)
Pediatrics	99 (13.9)	117 (15.1)	102 (12.2)
Emergency	50 (7.0)	60 (7.8)	62 (7.4)
Other	77 (10.7)	116 (15.0)	136 (16.3)
Age, mean (SD), y	27.5 (3.1)	27.6 (3.0)	27.5 (2.7)
PHQ-9 depression score, mean (SD), y	2.52 (3.11)	2.43 (3.05)	2.47 (3.07)
Neuroticism, mean (SD), y	21.4 (8.7)	22.0 (8.8)	21.4 (8.8)

Abbreviation: PHQ-9, Patient Health Questionnaire-9 depressive symptoms.

^aData are given as number (percentage) except where noted.

^bAll *P* values from 2-sample *t* tests (continuous variables) and χ^2 tests (categorical variables) comparing the 2009 and 2010 cohort with the 2011 cohort are nonsignificant (*P* > .05).

8099 baseline and quarterly participant assessments were completed.

INTERNSHIP COMPARISONS

Across the 4 quarterly assessments, the mean number of duty hours reported per week decreased significantly from a mean (SD) of 67.0 (17.0) in the preimplementation cohort to 64.3 (21.7) in the postimplementation cohort (*P* < .001). There was no significant effect of specialty on the change in work hours (*P* = .50). The percentage of interns who reported working more than 80 hours per week decreased from 12.8% to 7.8% (*P* < .001). In contrast, the mean number of reported hours slept each day was not significantly different between cohorts, with 6.8 (3.7) hours of sleep for the preimplementation cohort and 7.0 (4.3) hours of sleep for the postimplementation cohort (*P* = .17). The mean PHQ depression score during the year also did not change significantly between cohorts, with a mean PHQ depressive symptom score of 5.8 (4.8) for the preimplementation cohort and 5.7 (4.6) for the postimplementation cohort (*P* = .55). Correspondingly, the percentage of interns meeting PHQ criteria for depression (PHQ score \geq 10.0) did not change significantly between cohorts, with 20.0% of the preimplementation cohort and 18.7% of the postimplementation cohort meeting criteria for depression (*P* = .39). Consistent with the depressive symptom scores, the subjective well-being score did not change significantly under the new duty hour rules, with mean scores of 48.5 (13.9) for the preimplementation cohort and 48.4 (13.9) for the postimplementation cohort (*P* = .86). The percentage of interns who reported concern about making a serious medical error increased significantly from 19.9% in the preimplementation cohort to 23.3% in the postimplementation cohort (*P* = .007) (**Table 2** and **Figure**). There was no significant effect of specialty on the change in error rates (*P* = .15). Interns in both the preimplementation and postimplementation cohort who met PHQ criteria for depression were significantly more likely to report

a medical error compared with residents who did not meet criteria for depression (35.3% vs 17.8%; *P* < .001) (Table 2 and Figure).

COMMON PROGRAM SUBSAMPLE

Forty-two of the 51 residency programs participated in all 3 years of the study (N = 939 participants). In this subsample, the mean number of hours worked per week decreased from 67.1 (19.2) preimplementation to 63.5 (18.4) postimplementation. The magnitude of the decrease (3.6 hours) was similar to that in the full sample (2.7 hours) (*P* < .001). The percentage of interns who reported committing a serious medical error increased from 19.1% in the preimplementation cohort to 22.3% in the postimplementation cohort (*P* = .10). Although the increase in the percentage of interns committing a serious error with the implementation of the duty hour rules in the subsample (3.2%) was similar to that in the full sample (3.4%), the difference in the subsample did not reach statistical significance. All cohort differences in baseline and outcome measures that were nonsignificant in the overall sample were also nonsignificant in the subsample analyses. (Assessment of workload satisfaction and learning environment and for secular trends in error rates are included in the eAppendix; <http://www.jamainternalmed.com>.)

COMMENT

In this prospective, longitudinal, multi-institutional study, we found that during the 2011-2012 academic year, interns reported working fewer hours but more frequently reported concerns about committing medical errors than interns serving before the ACGME duty hour reforms were implemented. Furthermore, we found that the reforms were not associated with any reported changes in interns' sleep duration or their symptoms of depression.

Investigations of the 2003 ACGME duty hour restrictions differed as to whether residents worked fewer hours^{19,20} but generally agreed that these restrictions did not result in increased sleep for residents.^{19,21,22} With the more restrictive 2011 ACGME requirements, we found clear evidence that interns were working fewer hours. Similar to the 2003 reforms, however, our findings indicate that the 2011 ACGME restrictions did not significantly increase sleep for interns. In response to the new restriction on maximum shift length, many residency programs instituted a night float system. While we did not assess which programs switched to a night float system, our results are consistent with those of earlier work suggesting that the implementation of a night float system does not increase sleep for residents.²³ Some studies have suggested that changing away from traditional internship schedules can reduce the number of hours preceded by little or no sleep even without having a substantial effect on total sleep hours experienced by residents.^{12,24}

Previous work has demonstrated that physicians experience a substantial increase in depressive symptoms during internship.^{13,25} In this study, we found that fol-

Table 2. Duty Hours, Sleep, Depressive Symptoms, and Medical Errors in Pre-Duty Hour and Post-Duty Hour Restriction Cohorts

Variable	Quarter	2009	2010	2011
Duty hours per week ^b	First	68.8	67.1	66.7
	Second	68.4	65.9	63.8
	Third	68.1	65.7	64.7
	Fourth	67.1	65.4	61.5
Sleep, hours per day	First	6.9	7.1	7.2
	Second	6.8	6.5	7.1
	Third	6.9	6.4	7.0
	Fourth	6.8	6.5	6.5
PHQ-9 depressive symptom score	First	6.0	5.5	5.8
	Second	5.9	5.9	5.8
	Third	6.0	6.4	5.8
	Fourth	5.6	5.7	5.4
Serious medical error, % ^a	First	19.3	17.6	22.7
	Second	20.4	22.2	24.2
	Third	20.6	23.3	22.6
	Fourth	18.0	18.3	23.3

Abbreviation: PHQ-9, Patient Health Questionnaire-9 depressive symptoms.

^a $P < .05$.

^b $P < .001$.

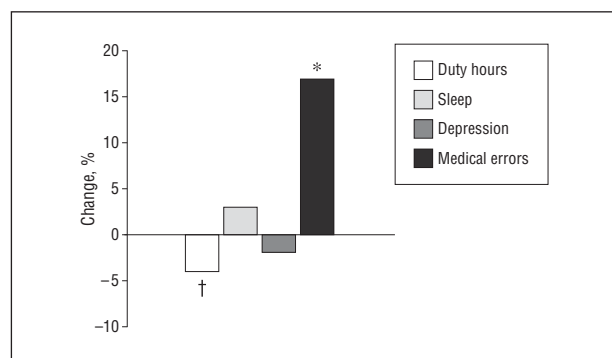


Figure. Percentage change in key variables before and after implementation of 2011 Accreditation Council for Graduate Medical Education (ACGME) reforms. Overall, work hours decreased by 4%, average number of hours of sleep per night over the past week increased by 3%, mean depressive symptoms decreased by 2%, and medical errors increase by 17% following the new ACGME guidelines. * $P < .05$; † $P < .001$.

Following implementation of the 2011 duty hour rules, the levels of depressive symptoms and well-being among interns were unchanged from those of previous cohorts. High levels of depression have been linked to more medical errors and poorer clinical performance.^{17,26} In this sample, we found further evidence supporting this link, with interns meeting PHQ criteria for depression reporting medical errors (35.3%) at almost twice the rate as non-depressed interns (17.8%).

Based on information from national mortality studies and single-institution longitudinal studies, the IOM concluded that the 2003 ACGME reforms did not harm patients and may have modestly improved outcomes for a subset of patients.¹ However, the magnitude of the improvement in patient safety was not as substantial as some had predicted, suggesting that a decrease in continuity of care may have largely offset a reduction in fatigue-related errors resulting from the reforms.¹ In contrast with the evidence suggesting a modest improvement in safety with the 2003 reforms, in this study of the 2011 reforms, we found unexpected evidence of an increase in

self-reported medical errors among interns. Although our analysis of errors in the common program subsample was not statistically significant, the difference in the rate of medical errors in the subsample was comparable with that in the overall sample, suggesting that the lack of significance may be due to inadequate statistical power.

There are multiple possible reasons why the 2011 ACGME reforms may not have achieved their stated goals. Given that increased sleep was a key mechanism through which the new duty hour restrictions were intended to improve the health of residents,³ the lack of such an effect in the postimplementation cohort in our study is a cause for concern. Designing work schedules that account for circadian phase and explicitly training residents on practices to increase sleep time and improve sleep quality may be necessary.²⁷

In addition, for many hospitals, the new duty hour restrictions were not accompanied by funding to hire additional clinical staff.²⁸ As a result, the duty hour restrictions may have exacerbated the problem of work compression, with residents expected to complete the same amount of work as previous cohorts but in less total time.²⁹ Increased work compression has been associated with poorer clinical performance and decreased satisfaction among residents.³⁰ Consistent with increased work compression, the 2011 duty hour reforms were associated with a modest, nonsignificantly lower level of workload satisfaction, despite being associated with a reduction in duty hours. Residents have also noted that the new duty hour rules necessitated a shift toward longer daily shifts and eliminated the reprieve of a “postcall” day.³¹ Evidence that resource-intensive interventions have been effective in improving resident satisfaction and patient care suggests that it may be necessary to allocate substantially more resources to achieve the improvements targeted by the ACGME reforms.³⁰

Finally, evidence strongly indicates that implementing the new ACGME reforms has resulted in an increased number of handoffs in most residency pro-

grams.³¹ The increase in handoffs may be a contributing factor to the increase in self-reported medical errors with the implementation of the new duty hour. While curricula in handoff training for residents have remained largely undeveloped, initial studies of standardized handoff systems have been effective in reducing errors.^{32,33}

There are important features of our study that strengthen confidence in the results. First, we studied large cohorts of residents, working in over a dozen different hospital systems, community and academic hospitals, and hospitals in all major regions of the United States. The full sample provided statistical power to detect effects of small to moderate size. Second, by focusing on interns, the target of the most aggressive and controversial aspects of the 2011 ACGME reforms, the study provides important information on the core aspects of the reforms. Third, the prospective design allows us to differentiate cohort outcomes from outcomes owing to pre-existing differences between the preimplementation and postimplementation cohorts. Fourth, by conducting quarterly assessments, we were able to obtain a picture of the ACGME reform effects throughout the year rather than a single cross-sectional snapshot.

In addition to these strengths there are limitations to our study. The most important limitation is the self-report nature of our assessment. Although self-report represents the only practical way to gather data on a large and varied sample of participants, this method is susceptible to participant bias, especially on topics as controversial as medical errors and duty hour reforms. We guarded against the influence of participant bias by embedding this study in a larger study with the stated aim of identifying genetic predictors of depression under stress, and, thus, the likelihood of participant bias in report of medical errors and duty hours was reduced.¹³ Furthermore, we did not ask interns about their opinions of duty hour reforms but queried them about more objective outcome measures of interest. Specifically related to medical errors, we did not find any evidence for trending in the self-reported error rates in the years preceding duty hour reforms or through the 4 quarters within years (eAppendix). This suggests that any long-term trends toward increasing patient safety awareness did not account for the increased rate of self-reported medical errors that we identified following the implementation of duty hour reforms. While the validity of self-report is questioned in skills assessment, self-report may be a more valid measure for variables assessed in this study.³⁴ For example, self-reported work hours match well with swipe-in and swipe-out electronic recordings of resident work hours.³⁵ Furthermore, some studies showed that most self-reported errors were confirmed by medical record review.^{36,37} Other work, however, has suggested that physician self-report and medical record review identified a similar number of errors, but that there was only partial overlap in the errors identified, with self-reported errors more likely to be preventable medical errors.³⁸ Although it is difficult to obtain gold standard measures (eg, through medical record review, electronic recording of work hours) on large samples across multiple institutions, investigators were able to conduct single-program objective assessment studies of the effect of the 2003 duty hour reforms.^{19,39} Similar studies evaluating the 2011

duty hour reforms would provide an important complement to this study.

There are other important limitations to our study. First, because the preimplementation cohort served in 2009 to 2010 and 2010 to 2011 and the postimplementation cohort served in 2011 to 2012, the differences between cohorts could have resulted from stricter enforcement of 2003 duty hour reforms, anticipation of future reforms, or secular trends unrelated to duty hour reforms. Because our study spanned 3 sequential academic years, however, it is unlikely that long-term trends could explain the observed effects. Furthermore, because the identified national trend has been toward improvements in the quality of care, temporal trends would, if anything, have served to obscure effects of duty hour changes.⁴⁰ Second, only 58% of invited individuals chose to participate in the study. Although this is a comparatively high participation rate for a multi-institutional study, any systematic differences between participants and non-participants could introduce bias. Finally, our study assessed only the effects of duty hour reforms during the first year of their implementation. Studies should assess changes in interns' sleep and rates of depression and medical errors in future years, after hospital systems have had time to adjust to the new duty hour restrictions.

In conclusion, this large prospective longitudinal study of the 2011 ACGME duty hour reforms found that, although the reforms reduced the total number of hours that interns are on duty, they did not affect interns' duration of sleep or mental health and increased the frequency of self-reported medical errors. Different strategies for improving resident education and patient care may be necessary to achieve the desired impact of ACGME reforms.

Accepted for Publication: January 1, 2013.

Published Online: March 25, 2013. doi:10.1001/jamainternmed.2013.351

Author Affiliations: Departments of Psychiatry (Drs Sen and Dalack) and Internal Medicine (Dr Kolars), University of Michigan, Ann Arbor; Department of Psychiatry, Perelman School of Medicine, and University of Pennsylvania, Philadelphia VA Medical Center, Philadelphia (Dr Kranzler); Department of Internal Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois (Dr Didwania); Department of Psychiatry, Emory University School of Medicine, Atlanta, Georgia (Dr Schwartz); Department of Radiation Oncology, University of Washington, Seattle, (Dr Amarnath); Departments of Pediatrics and Medicine, Keck University of Southern California School of Medicine, Los Angeles County + University of Southern California Medical Center, Los Angeles (Dr Nichols); and Department of Psychiatry, Medical University of South Carolina, Charleston (Dr Guille).

Correspondence: Srijan Sen, MD, PhD, Department of Psychiatry, University of Michigan, 5047 BSRB, 109 Zina Pitcher Pl, Ann Arbor, MI 48109 (srijan@umich.edu).

Author Contributions: Dr Sen had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Sen, Schwartz, and Guille. *Acquisition of data:* Sen, Nichols, and Guille. *Analysis and inter-*

pretation of data: Sen, Kranzler, Didwania, Amarnath, Kolars, Dalack, and Guille. *Drafting of the manuscript*: Sen, and Kolars. *Critical revision of the manuscript for important intellectual content*: Sen, Kranzler, Didwania, Schwartz, Amarnath, Kolars, Dalack, Nichols, and Guille. *Statistical analysis*: Sen. *Obtained funding*: Sen and Guille. *Administrative, technical, and material support*: Sen, Kranzler, Didwania, Schwartz, Dalack, Nichols, and Guille. *Study supervision*: Sen.

Conflict of Interest Disclosures: Dr Krantzler has been a consultant or advisory board member for Alkermes, Lilly, Lundbeck, Pfizer, and Roche. He is also a member of the American Society of Clinical Psychopharmacology, which is supported by Lilly, Lundbeck, Abbott, and Pfizer.

Funding/Support: The project was supported by the following grants: a Young Investigator Grant from the American Foundation for Suicide Prevention (Dr Sen), UL1RR024986 from the National Center for Research Resources (Dr Sen), MH095109 from the National Institute of Mental Health (Dr Sen), and AA013736 from the National Institute on Alcohol Abuse and Alcoholism (Dr Krantzler).

Role of the Sponsors: The funding agencies played no role in the design and conduct of the study; collection management, analysis, or interpretation of the data; and preparation, review, or approval of the manuscript.

Online-Only Material: The eAppendix is available at <http://www.jamainternalmed.com>.

Additional Contributions: Faren Grant, BA (University of Michigan), and Heather Bryant, MA (University of Michigan), served as study coordinators for the project and were compensated for their work on the Intern Health Study. We thank the participating interns and program directors for the time that they invested in this study.

REFERENCES

1. Ulmer C, Wolman DM, Johns MME. *Resident Duty Hours: Enhancing Sleep, Supervision, and Safety*. Washington, DC: National Academies Press; 2009.
2. Kuehn BM. New rules call for more oversight, fewer hours for first-year residents. *JAMA*. 2010;304(9):950-951.
3. Nasca TJ, Day SH, Amis ES Jr; ACGME Duty Hour Task Force. The new recommendations on duty hours from the ACGME Task Force. *N Engl J Med*. 2010; 363(2):e3.
4. 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):1838-1848.
5. Hart RP, Buchsbaum DG, Wade JB, Hamer RM, Kwentus JA. Effect of sleep deprivation on first-year residents' response times, memory, and mood. *J Med Educ*. 1987;62(11):940-942.
6. Levine AC, Adusumilli J, Landrigan CP. Effects of reducing or eliminating resident work shifts over 16 hours: a systematic review. *Sleep*. 2010;33(8):1043-1053.
7. Barger LK, Cade BE, Ayan NT, et al; Harvard Work Hours, Health, and Safety Group. Extended work shifts and the risk of motor vehicle crashes among interns. *N Engl J Med*. 2005;352(2):125-134.
8. Valko RJ, Clayton PJ. Depression in the internship. *Dis Nerv Syst*. 1975;36(1):26-29.
9. Iglehart JK. The ACGME's final duty-hour standards: special PGY-1 limits and strategic napping. *N Engl J Med*. 2010;363(17):1589-1591.
10. Vidyarthi AR, Arora V, Schnipper JL, Wall SD, Wachter RM. Managing discontinuity in academic medical centers: strategies for a safe and effective resident sign-out. *J Hosp Med*. 2006;1(4):257-266.
11. Petersen LA, Brennan TA, O'Neil AC, Cook EF, Lee TH. Does housestaff discontinuity of care increase the risk for preventable adverse events? *Ann Intern Med*. 1994;121(11):866-872.
12. Auger KA, Landrigan CP, Gonzalez del Rey JA, Sieplinga KR, Sucharew HJ, Simons JM. Better rested, but more stressed? evidence of the effects of resident work hour restrictions. *Acad Pediatr*. 2012;12(4):335-343.
13. Sen S, Kranzler HR, Krystal JH, et al. A prospective cohort study investigating factors associated with depression during medical internship. *Arch Gen Psychiatry*. 2010;67(6):557-565.
14. Costa PT Jr, McCrae RR. Stability and change in personality assessment: the revised NEO Personality Inventory in the year 2000. *J Pers Assess*. 1997;68(1):86-94.
15. Costa PT, McCrae RR. *The NEO Personality Inventory*. Odessa, FL: Psychological Assessment Resources; 1985.
16. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613.
17. West CP, Huschka MM, Novotny PJ, et al. Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study. *JAMA*. 2006;296(9):1071-1078.
18. Keyes CL. Mental illness and/or mental health? investigating axioms of the complete state model of health. *J Consult Clin Psychol*. 2005;73(3):539-548.
19. Landrigan CP, Fahrenkopf AM, Lewin D, et al. Effects of the accreditation council for graduate medical education duty hour limits on sleep, work hours, and safety. *Pediatrics*. 2008;122(2):250-258.
20. Jaggi R, Weinstein DF, Shapiro J, Kitch BT, Dorer D, Weissman JS. The Accreditation Council for Graduate Medical Education's limits on residents' work hours and patient safety: a study of resident experiences and perceptions before and after hours reductions. *Arch Intern Med*. 2008;168(5):493-500.
21. Arora VM, Georgitis E, Siddique J, et al. Association of workload of on-call medical interns with on-call sleep duration, shift duration, and participation in educational activities. *JAMA*. 2008;300(10):1146-1153.
22. Fitzgibbons SC, Chen J, Jaggi R, Weinstein D. Long-term follow-up on the educational impact of ACGME duty hour limits: a pre-post survey study. *Ann Surg*. 2012;256(6):1108-1112.
23. Chua KP, Gordon MB, Sectish T, Landrigan CP. Effects of a night-team system on resident sleep and work hours. *Pediatrics*. 2011;128(6):1142-1147.
24. Lockley SW, Cronin JW, Evans EE, et al; Harvard Work Hours, Health and Safety Group. Effect of reducing interns' weekly work hours on sleep and attentional failures. *N Engl J Med*. 2004;351(18):1829-1837.
25. Bellini LM, Baime M, Shea JA. Variation of mood and empathy during internship. *JAMA*. 2002;287(23):3143-3146.
26. Fahrenkopf AM, Sectish TC, Barger LK, et al. Rates of medication errors among depressed and burnt out residents: prospective cohort study. *BMJ*. 2008;336(7642):488-491.
27. Arora VM, Georgitis E, Woodruff JN, Humphrey HJ, Meltzer D. Improving sleep hygiene of medical interns: can the sleep, alertness, and fatigue education in residency program help? *Arch Intern Med*. 2007;167(16):1738-1744.
28. Romano PS, Volpp K. The ACGME's 2011 changes to resident duty hours: are they an unfunded mandate on teaching hospitals? *J Gen Intern Med*. 2012; 27(2):136-138.
29. Ludmerer KM. Redesigning residency education: moving beyond work hours. *N Engl J Med*. 2010;362(14):1337-1338.
30. McMahon GT, Katz JT, Thorndike ME, Levy BD, Loscalzo J. Evaluation of a redesign initiative in an internal-medicine residency. *N Engl J Med*. 2010;362(14):1304-1311.
31. Drolet BC, Christopher DA, Fischer SA. Residents' response to duty-hour regulations: a follow-up national survey. *N Engl J Med*. 2012;366(24):e35.
32. Okie S. An elusive balance: residents' work hours and the continuity of care. *N Engl J Med*. 2007;356(26):2665-2667.
33. Petersen LA, Orav EJ, Teich JM, O'Neil AC, Brennan TA. Using a computerized sign-out program to improve continuity of inpatient care and prevent adverse events. *Jt Comm J Qual Improv*. 1998;24(2):77-87.
34. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol*. 1999;77(6):1121-1134.
35. Todd SR, Fahy BN, Paukert JL, Mersinger D, Johnson ML, Bass BL. How accurate are self-reported resident duty hours? *J Surg Educ*. 2010;67(2):103-107.
36. Weingart SN, Callanan LD, Ship AN, Aronson MD. A physician-based voluntary reporting system for adverse events and medical errors. *J Gen Intern Med*. 2001; 16(12):809-814.
37. November M, Chie L, Weingart SN. Physician-reported adverse events and medical errors in obstetrics and gynecology. In: Henriksen K, Battles JB, Keyes MA, Grady ML, eds. *Advances in Patient Safety: New Directions and Alternative Approaches*. Vol 1. Rockville, MD: Assessment; 2008.
38. O'Neil AC, Petersen LA, Cook EF, Bates DW, Lee TH, Brennan TA. Physician reporting compared with medical-record review to identify adverse medical events. *Ann Intern Med*. 1993;119(5):370-376.
39. 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):984-992.
40. Horwitz LI, Kosiborod M, Lin Z, Krumholz HM. Changes in outcomes for internal medicine inpatients after work-hour regulations. *Ann Intern Med*. 2007;147(2):97-103.