



Published in final edited form as:

JAMA Intern Med. 2013 April 22; 173(8): 657–663. doi:10.1001/jamainternmed.2013.351.

Effects of the 2011 Duty Hour Reforms on Interns and Their Patients: A Prospective Longitudinal Cohort Study

Dr. Srijan Sen, M.D. Ph.D.,

Department of Psychiatry, University of Michigan

Dr. Henry R. Kranzler, M.D.,

Department of Psychiatry, Perelman School of Medicine of the University of Pennsylvania and the Philadelphia VAMC

Dr. Aashish K. Didwania, M.D.,

Department of Internal Medicine, Northwestern University Feinberg School of Medicine

Dr. Ann C. Schwartz, M.D.,

Department of Psychiatry, Emory University School of Medicine

Dr. Sudha Amarnath, M.D.,

Department of Radiation Oncology, University of Washington

Dr. Joseph C. Kolars, M.D.,

Department of Internal Medicine, University of Michigan

Dr. Gregory W. Dalack, M.D.,

Department of Internal Medicine, University of Michigan

Dr. Breck Nichols, M.D. M.P.H, and

Departments of Pediatrics and Medicine, Keck USC School of Medicine/LAC+USC Medical Center

Dr. Constance Guille, M.D.

Department of Psychiatry, Medical University of South Carolina

Abstract

Background—In 2003, the first phase of duty hour requirements for U.S. 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 wellbeing 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. We sought to determine the effects of the 2011 duty hour reforms on first year residents (interns) and their patients.

Location of work and address for correspondence: Srijan Sen, Department of Psychiatry, University of Michigan, 5047 BSRB, 109 Zina Pitcher Place, Ann Arbor, MI 48109, Phone:(734) 395-8319, Fax:(734) 936-2690, srijan@umich.edu.

Financial Disclosures: None

Methods—We conducted a longitudinal cohort study of 2323 interns entering one of 51 residency programs at 14 university and community-based GME institutions or graduating from one of four medical schools participating in the study. We compared 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 between interns serving before (2009 and 2010) and interns serving after (2011) the implementation of the new duty-hour requirements.

Results—58% of invited interns chose to participate in the study. Reported duty hours decreased from an average of 67.0 hours/week before the new rules to 64.3 hours/week after the new rules were instituted ($p<0.001$). Despite the decrease in duty hours, there were no significant changes in hours slept (7.0→6.8; $p=0.17$), depressive symptoms (5.8→5.7; $p=NS$) or well-being (48.5→48.4; $p=0.86$) reported by interns. With the new duty hour rules, the percentage of interns who reported committing a serious medical error increased from 19.9% to 23.3% ($p=0.007$).

Conclusions—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 wellbeing but has been accompanied by an unanticipated increase in self-reported medical errors under the new duty hour restrictions.

Keywords

Graduate; Medical; Education; Residency; Work; Hours; Sleep

Background

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 Recognition of this concern, has sparked intense debate among policy makers, medical educators, and residents themselves about how to devise a medical education system that promotes rigorous, high-quality training while maintaining the health of residents and maximizing the safety of patients.²

Using input from the Institute of Medicine (IOM), the Accreditation Council for Graduate Medical Education (ACGME) established a new set of duty hour recommendations that took effect July, 2011.^{1,3} The ACGME recommendations focus on establishing progressive responsibility through residency, with greater oversight and lower work demand for first-year residents (commonly referred to as interns), because they are the least experienced clinicians and traditionally have worked the longest hours. While the recommendations included numerous changes in supervision and in the requirement for greater oversight by programs, the most important and controversial changes related to maximum shift length for interns. A landmark study published in 2004 found that an intensive intervention designed to reduce extended intern shifts in the ICU successfully reduced the incidence of serious medical errors.⁴ This study, along with additional work indicating that long duty hours harms both patients^{5,6} and residents themselves,^{7,8} prompted the ACGME to restrict the maximum shift length for interns to 16 hours.⁹ Other research, however, has shown a substantial increase in patient handoffs with shift length restriction and a higher rate of

medical errors with more handoffs.^{10,11} Further, a recent study found that duty hour restrictions increased the stress felt by residents, despite a reduction in extended shifts.¹² As a result, some residency program directors have warned that the new ACGME restrictions would result in poorer training for interns and poorer outcomes for patients.^{13,14} These conflicting points of view highlight the importance of evaluating the effect of the duty hour changes for interns and their patients.

As part of the ongoing Intern Health Study,¹⁵ we conducted a large, prospective, longitudinal study comparing interns serving before and interns serving after the new duty hour standards were instituted. The study solicited self-report information to evaluate the effects of the new ACGME duty hour standards on resident duty hours, sleep duration, depressive symptoms and medical errors in a sample of interns from multiple specialties and institutions across the country.

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.¹⁵ Fifty-one residency programs at 10 university-based and four community-based GME institutions agreed to allow us to invite their incoming residents to take part in the study. In addition, four medical schools agreed to allow us to invite their graduating students to take part in the study. In total, 4352 individuals entering internal medicine, general surgery, pediatrics, obstetrics/gynecology, emergency medicine, combined medicine/pediatrics, transitional year and psychiatry residency programs during the 2009, 2010 and 2011 academic years were sent an email two months prior to commencing internship and invited to participate in the study.

For 348 subjects, our email invitations were returned as undeliverable and we were unable to obtain a valid email address. Fifty-eight percent (2323/4005) of the remaining invited subjects agreed to participate in the study. The Institutional Review Board at all of the participating institutions approved the study. Subjects were compensated \$40 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 subjects at residency programs that took part in the study in all three years of study (See Supplementary Online Content for “Common program sub-sample analysis”).

Data Collection

The data collection procedures utilized in the Intern Health Study have been detailed previously.¹⁵ Briefly, all surveys were conducted through a secure online website designed to maintain confidentiality, with subjects identified only by numbers.

Pre-Internship Assessment—Subjects 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 depressive symptoms (PHQ-9), history of depression and a measure of personality, the NEO-Five Factor

Inventory Neuroticism).^{16,17} The PHQ-9 is a nine-item self-report component of the PRIME-MD inventory, designed to screen for depressive symptoms. A score of 10 or greater on the PHQ-9 has a sensitivity of 93% and a specificity of 88% for the diagnosis of major depressive disorder.¹⁸

Within-Internship Assessments—Interns were contacted via email on the first Tuesday of months 3, 6, 9 and 12 of internship year and asked to complete a previously published survey addressing work hours, sleep and medical error.¹⁵ 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?”¹⁹ 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 6.²⁰

Statistical Analysis

All analyses were performed using SPSS 19.0 (Chicago, IL). Baseline factors were compared between the pre-implementation and post-implementation cohorts with a two-sample t-test for continuous measures and a Chi-Squared test for categorical measures. A series of Generalized Estimating Equation (GEE) analyses were utilized to assess the effects of the duty hour rule, while accounting for the repeated measures within subjects during internship. In the GEE analyses, cohort membership was entered as the predictor variable (pre-implementation (2009 and 2010) cohort compared to post-implementation (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 × specialty interaction term was included in these analyses. Two-sample t-tests were used to compare Subjective Well-Being between cohorts.

Results

Representativeness of the Sample

Compared to 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 chose to participate in the current study and individuals who chose not to participate (demographic information for full residency programs provided by American Association of Medical Colleges; personal communication - September 17, 2012).

Pre-Internship Comparisons

For the 2323 study participants, there were no significant differences in specialty, baseline depressive symptom score, neuroticism, gender or age between the pre-implementation (2009 (N=714) and 2010 (N=772)) and post-implementation (2011 (N=837)) cohorts (Table 1; all p values=NS). Further, there were no differences in these baseline characteristics

between subjects who completed one, two, three or four quarterly surveys (all p values=NS). On average, 65% of participants responded to each quarterly survey. In total, 8,099 baseline and quarterly subject assessments were completed.

Internship Comparisons

Across the four quarterly assessments, the mean number of duty hours reported per week decreased significantly from a mean (SD) of 67.0 ± 17.0 in the pre-implementation cohort to 64.3 ± 21.7 in the post-implementation cohort (Wald Chi Square(1)=31.6; $p < 0.001$). There was no significant effect of specialty on the change in work hours (Wald Chi Square(4)=3.34; $p = 0.50$). The percentage of interns who reported working more than 80 hours per week decreased from 12.8% to 7.8% (Wald Chi Square(1)=28.1; $p < 0.001$). 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 pre-implementation cohort and 7.0 ± 4.3 hours of sleep for the post-implementation cohort (Wald Chi Square(1)=2.3; $p = 0.17$). 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 pre-implementation cohort and 5.7 ± 4.6 for the post-implementation cohort (Wald Chi Square(1)=0.37; $p = 0.55$). Correspondingly, the percentage of interns meeting PHQ criteria for depression (PHQ score ≥ 10) did not change significantly between cohorts, with 20.0% of the pre-implementation cohort and 18.7% of the post-implementation cohort meeting criteria for depression (Wald Chi Square(1)=0.75; $p = 0.39$). 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 pre-implementation cohort and 48.4 ± 13.9 for the post-implementation cohort ($t(1) = -0.17$; $p = 0.86$). The percentage of interns who reported concern about making a serious medical error increased significantly from 19.9% in the pre-implementation cohort to 23.3% in the post-implementation cohort (Wald Chi Square(1)=7.3; $p = 0.007$) (Table 2; Figure 1). There was no significant effect of specialty on the change in error rates (Wald Chi Square(4)=6.74; $p = 0.15$). Interns in both the pre-implementation and post-implementation cohort who met PHQ criteria for depression were significantly more likely to report a medical error compared to residents who did not meet criteria for depression (35.3% vs. 17.8%; Wald Chi Square(1)=92.3; $p < 0.001$) (Table 2; Figure 1). Results from Common program sub-sample analysis are presented in the Supplementary Online Materials.

Discussion

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. Further, we find 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 yielded conflicting findings on whether they resulted in residents working fewer hours,^{21,22} but generally agreed that these restrictions did not result in increased sleep for residents.^{21,23,24} 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 in 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 earlier work suggesting that the implementation of a night-float system does not increase sleep for residents.²⁵ 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^{12,26}. Future studies should investigate whether the timing or quality of resident sleep or the number of hours worked on little or no sleep has changed with the new restrictions.

Previous work has demonstrated that physicians experience a substantial increase in depressive symptoms during internship.^{15,27} In this study, we found that the levels of depressive symptoms and well-being among interns were unchanged from previous cohorts, following implementation of the 2011 duty hour rules. High levels of depression have been linked to more medical errors and poorer clinical performance.^{19,28} In this sample, we find 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 pre-post 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 to 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. Specifically, the average percentage of residents reporting a serious medical error during each quarter increased from 19.9% before implementation of the 2011 ACGME reforms to 23.3% after implementation. Although our analysis of errors in the common program sub-sample did not reach significance, the difference in the rate of medical errors in the sub-sample was comparable to that in the overall sample, suggesting that the lack of significance may be due to inadequate statistical power (see Supplementary Online Content).

There are multiple possible reasons that the 2011 ACGME reforms may have yet to achieve their stated goals. Given that increased sleep was a key mechanisms through which the new duty hour restrictions were intended to improve the health of residents,³ the lack of such an effect in the post-implementation cohort in our study is concerning. 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.²⁹

In addition, for many hospitals, the new duty hour restrictions were not accompanied by funding to hire additional clinical staff.³⁰ 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.³¹ Increased work compression has been associated with poorer clinical performance and decreased satisfaction among residents.³² Consistent with increased work compression, the 2011 duty hour reforms were associated with a modest, non-significantly 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 towards longer daily shifts and eliminated the reprieve of a “post-call” day.³³ 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.³²

Finally, evidence strongly indicates that implementing the new ACGME reforms has resulted in an increased number of handoffs in most residency programs.³³ 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 has remained largely undeveloped, initial studies of standardized handoff systems have been effective in reducing errors.^{34,35} Thus, more research in handoff training and broader implementation of standardized systems to improve handoffs may be necessary to ensure high quality patient care under the new duty hour regulations.³⁶

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 due to pre-existing differences between the pre- and post-implementation 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 subjects, this method is susceptible to subject bias, especially on topics as controversial as medical errors and duty hour reforms. We guarded against the influence of subject 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 subject bias in report of medical errors and duty hours was reduced.¹⁵ Further, we did not ask interns about their opinions on 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 four quarters within years (see Supplementary Online Content). This suggests that any long-term trends towards 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.³⁷ For example, self-reported work hours matches well with swipe-in and swipe-out electronic recordings of resident work hours.³⁸ Further, some studies showed that the substantial majority of self-reported errors were confirmed by medical record review.^{39,40} 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 errors⁴¹. Although it is difficult to obtain gold standard measures (e.g., through medical record review, electronic recording of work hours, etc.) 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.^{21,42} Going forward, similar studies evaluating the 2011 duty hour reforms would provide a vitally important complement to the multi-institutional self-report results reported here.

There are other important limitations to our study. First, because the pre-implementation cohort served in 2009-10 and 2010-11 and the post-implementation cohort served in 2011-12, 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 three sequential academic years, however, it is unlikely that long-term trends could explain the observed effects. Further, because the identified national trend has been towards improvements in the quality of care, temporal trends would, if anything, have served to obscure effects of duty hour changes.⁴³ 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. Despite the lack of demographic differences between individuals who chose to take part in the study and those who did not, our results should be extrapolated with caution. Finally, our study only assessed 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 pre-post study of the 2011 ACGME duty hour reforms and 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. These concerning findings represent important early evidence that the newest ACGME restrictions, alone, have not met with immediate success in either improving the health of residents, or in leading to a reduction in self-reported errors. Further changes, targeted specifically towards improving resident education and patient care, may be necessary to achieve the desired impact of ACGME reforms.

Acknowledgments

We thank Faren Grant B.A. (University of Michigan) and Heather Bryant M.A (University of Michigan) who served as study coordinators for the project. Most importantly, we thank the participating interns and program directors for the time that they invested in this study. The project was supported by the following grants: *UL1RR024986* from the National Center for Research Resources (SS), *MH095109* from the National Institute of Mental Health (SS), *AA013736* from the National Institute on Alcohol Abuse and Alcoholism (HK), and a Young Investigator Grant from the American Foundation for Suicide Prevention (SS). 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. 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.

References

1. Ulmer, C.; Wolman, DM.; Johns, MME. Resident Duty Hours: Enhancing Sleep, Supervision, and Safety. Washington, DC: The National Academies Press; 2009.
2. Kuehn BM. New rules call for more oversight, fewer hours for first-year residents. *JAMA: the journal of the American Medical Association*. Sep 1; 2010 304(9):950–951. [PubMed: 20810365]
3. Nasca TJ, Day SH, Amis ES Jr. The new recommendations on duty hours from the ACGME Task Force. *N Engl J Med*. Jul 8.2010 363(2):e3. [PubMed: 20573917]
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*. Oct 28; 2004 351(18):1838–1848. [PubMed: 15509817]
5. Hart RP, Buchsbaum DG, Wade JB, Hamer RM, Kwentus JA. Effect of sleep deprivation on first-year residents' response times, memory, and mood. *Journal of medical education*. Nov; 1987 62(11): 940–942. [PubMed: 3681926]
6. Levine AC, Adusumilli J, Landrigan CP. Effects of reducing or eliminating resident work shifts over 16 hours: a systematic review. *Sleep*. Aug; 2010 33(8):1043–1053. [PubMed: 20815185]
7. Barger LK, Cade BE, Ayas NT, et al. Extended work shifts and the risk of motor vehicle crashes among interns. *N Engl J Med*. Jan 13; 2005 352(2):125–134. [PubMed: 15647575]
8. Valko RJ, Clayton PJ. Depression in the internship. *Dis Nerv Syst*. Jan; 1975 36(1):26–29. [PubMed: 1109883]
9. Iglehart JK. The ACGME's final duty-hour standards-special PGY-1 limits and strategic napping. *N Engl J Med*. Oct 21; 2010 363(17):1589–1591. [PubMed: 20973151]
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. *Journal of hospital medicine: an official publication of the Society of Hospital Medicine*. Jul; 2006 1(4):257–266. [PubMed: 17219508]
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*. Dec 1; 1994 121(11):866–872. [PubMed: 7978700]
12. Auger KA, Landrigan CP, Gonzalez Del, Rey JA, Sieplinga KR, Sucharew HJ, Simmons JM. Better Rested, but More Stressed? Evidence of the Effects of Resident Work Hour Restrictions. *Acad Pediatr*. 2012 May 26.12:335–343. [PubMed: 22626586]
13. Rice LB. Evidence-based evaluation of physician work hour regulations. *JAMA: the journal of the American Medical Association*. Feb 4.2009 301(5):484. author reply 484-485. [PubMed: 19190310]
14. Shea JA, Willett LL, Borman KR, et al. Anticipated Consequences of the 2011 Duty Hours Standards: Views of Internal Medicine and Surgery Program Directors. *Acad Med*. Jul; 2012 87(7):895–903. [PubMed: 22622221]
15. Sen S, Kranzler HR, Krystal JH, et al. A prospective cohort study investigating factors associated with depression during medical internship. *Arch Gen Psychiatry*. Jun; 2010 67(6):557–565. [PubMed: 20368500]
16. 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. [PubMed: 9018844]
17. Costa, PaM; R, R. The NEO Personality Inventory: Manual. 1985
18. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. Sep; 2001 16(9):606–613. [PubMed: 11556941]
19. West CP, Huschka MM, Novotny PJ, et al. Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study. *JAMA*. Sep 6; 2006 296(9):1071–1078. [PubMed: 16954486]

20. Keyes CL. Mental illness and/or mental health? Investigating axioms of the complete state model of health. *Journal of consulting and clinical psychology*. Jun; 2005 73(3):539–548. [PubMed: 15982151]
21. 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*. Aug; 2008 122(2): 250–258. [PubMed: 18676540]
22. 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*. Mar 10; 2008 168(5):493–500. [PubMed: 18332295]
23. 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*. Sep 10; 2008 300(10):1146–1153. [PubMed: 18780843]
24. 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*. Oct 12.2012
25. Chua KP, Gordon MB, Sectish T, Landrigan CP. Effects of a night-team system on resident sleep and work hours. *Pediatrics*. Dec; 2011 128(6):1142–1147. [PubMed: 22123867]
26. Lockley SW, Cronin JW, Evans EE, et al. Effect of reducing interns' weekly work hours on sleep and attentional failures. *The New England journal of medicine*. Oct 28; 2004 351(18):1829–1837. [PubMed: 15509816]
27. Bellini LM, Baime M, Shea JA. Variation of mood and empathy during internship. *JAMA*. Jun 19; 2002 287(23):3143–3146. [PubMed: 12069680]
28. Fahrenkopf AM, Sectish TC, Barger LK, et al. Rates of medication errors among depressed and burnt out residents: prospective cohort study. *BMJ*. Mar 1; 2008 336(7642):488–491. [PubMed: 18258931]
29. 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*. Sep 10; 2007 167(16):1738–1744. [PubMed: 17846392]
30. 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*. Feb; 2012 27(2):136–138. [PubMed: 22127799]
31. Ludmerer KM. Redesigning residency education--moving beyond work hours. *N Engl J Med*. Apr 8; 2010 362(14):1337–1338. [PubMed: 20375412]
32. McMahon GT, Katz JT, Thorndike ME, Levy BD, Loscalzo J. Evaluation of a redesign initiative in an internal-medicine residency. *N Engl J Med*. Apr 8; 2010 362(14):1304–1311. [PubMed: 20375407]
33. Drolet BC, Christopher DA, Fischer SA. Residents' response to duty-hour regulations--a follow-up national survey. *The New England journal of medicine*. Jun 14.2012 366(24):e35. [PubMed: 22646511]
34. Okie S. An elusive balance--residents' work hours and the continuity of care. *The New England journal of medicine*. Jun 28; 2007 356(26):2665–2667. [PubMed: 17596598]
35. 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. *The Joint Commission journal on quality improvement*. Feb; 1998 24(2):77–87. [PubMed: 9547682]
36. Wohlauer MV, Arora VM, Horwitz LI, Bass EJ, Mahar SE, Philibert I. The patient handoff: a comprehensive curricular blueprint for resident education to improve continuity of care. *Acad Med*. Apr; 2012 87(4):411–418. [PubMed: 22361791]
37. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*. Dec; 1999 77(6):1121–1134. [PubMed: 10626367]
38. Todd SR, Fahy BN, Paukert JL, Mersinger D, Johnson ML, Bass BL. How accurate are self-reported resident duty hours? *J Surg Educ*. Mar-Apr;2010 67(2):103–107. [PubMed: 20656607]
39. Weingart SN, Callanan LD, Ship AN, Aronson MD. A physician-based voluntary reporting system for adverse events and medical errors. *J Gen Intern Med*. Dec; 2001 16(12):809–814. [PubMed: 11903759]

40. 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., editors. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol 1: Assessment)*. Rockville (MD): 2008.
41. 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*. Sep 1; 1993 119(5):370–376. [PubMed: 8338290]
42. 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*. Sep 5; 2007 298(9):984–992. [PubMed: 17785643]
43. Horwitz LI, Kosiborod M, Lin Z, Krumholz HM. Changes in outcomes for internal medicine inpatients after work-hour regulations. *Ann Intern Med*. Jul 17; 2007 147(2):97–103. [PubMed: 17548401]

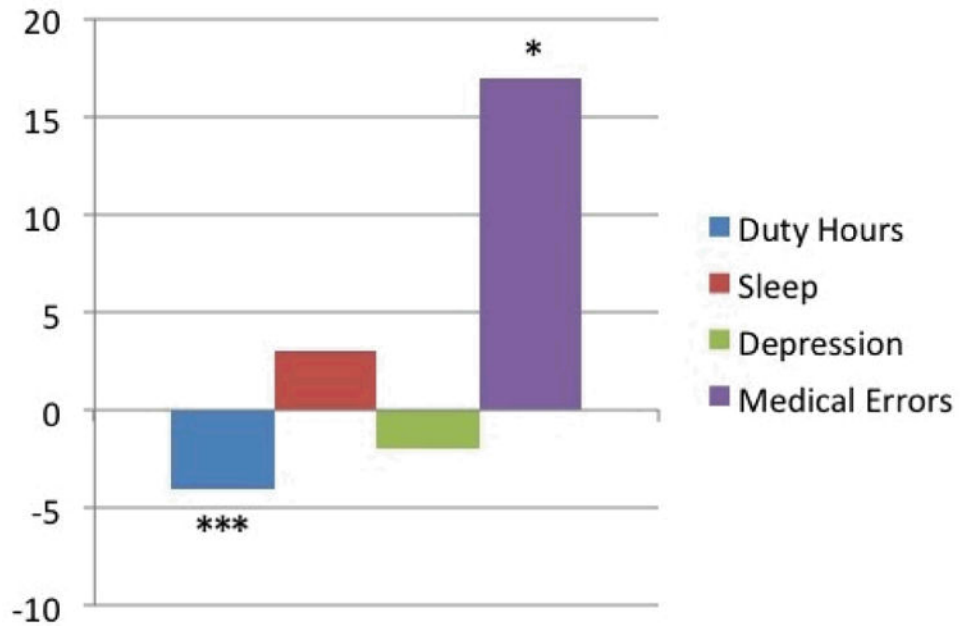


Figure 1. Percent Change in Key Variables Pre and Post Implementation of 2011 ACGME Reforms

Figure 1 shows the percent change in outcomes pre and post 2011 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. *indicates $P < 0.05$; *** indicates $P < 0.001$.

Table 1

Sample Demographic Characteristics (N(%))

Variable	2009	2010	2011	p value#	
Sex	Male	345 (48.3)	383 (49.6)	412 (49.2)	NS
	Female	369 (51.7)	389 (50.4)	425 (50.8)	
Specialty	Internal Medicine	404 (56.6)	412 (53.4)	451 (53.9)	NS
	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	27.5 (3.1)	27.6 (3.0)	27.5 (2.7)	NS	
PHQ Depression Score	2.52 (3.11)	2.43 (3.05)	2.47 (3.07)	NS	
Neuroticism	21.4 (8.7)	22.0 (8.8)	21.4 (8.8)	NS	

- Two sample T-test (continuous variables) or Chi Square (categorical variables) p value comparing 2009+2010 cohort to 2011 cohort

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

		2009	2010	2011
Duty Hours (hours/week)***	1 st quarter	68.76	67.07	66.68
	2 nd quarter	68.40	65.87	63.78
	3 rd quarter	68.08	65.70	64.65
	4 th quarter	67.05	65.41	61.45
Sleep (hours/day)	1 st quarter	6.85	7.13	7.24
	2 nd quarter	6.84	6.51	7.08
	3 rd quarter	6.87	6.35	6.99
	4 th quarter	6.82	6.48	6.52
PHQ Depressive Symptom Score	1 st quarter	6.02	5.47	5.77
	2 nd quarter	5.87	5.86	5.80
	3 rd quarter	5.99	6.35	5.84
	4 th quarter	5.56	5.67	5.42
Serious Medical Error (%)*	1 st quarter	19.3%	17.6%	22.7%
	2 nd quarter	20.4%	22.2%	24.2%
	3 rd quarter	20.6%	23.3%	22.6%
	4 th quarter	18.0%	18.3%	23.3%

* indicates $P < 0.05$;

*** indicates $P < 0.001$