Final Progress Report

Title: Effects of Extended Work Hours on ICU Patient Safety

Principal Investigator: Charles A. Czeisler, Ph.D., M.D.

Organization: Brigham and Women's Hospital

Co-Investigators: Christopher P. Landrigan, M.D., M.P.H., Jeffrey M. Rothschild, M.D., M.P.H., John W. Cronin, M.D., Rainu Kaushal, M.D., M.P.H., Elisabeth Burdick, M.S., Joel T. Katz, M.D., Craig M. Lilly, M.D., Peter H. Stone, M.D., Steven W. Lockley, Ph.D., Erin Evans, B.A., David W. Bates, M.D., M.Sc., Orfeu Buxton, Ph.D., Brian E. Cade, M.S.

Inclusive Dates: 09/30/01-09/29/04

Federal Project Officer: Eileen Hogan

Agency: Agency for Healthcare Research and Quality

Grant number: RO1 HS012032

Abstract

Purpose: The goals of this study were to compare work hours, sleep, attentional failures, and medical errors among interns working a traditional schedule and an intervention schedule that eliminated traditional 30-hour-in-a-row shifts.

Scope: Few systematic studies of sleep deprivation in clinical settings have been conducted. We sought to conduct a randomized experiment to test the effect on patient safety of implementing shorter work shifts for interns in ICU settings.

Methods: We compared interns' sleep, work, and alertness as well as ICU patient safety in two intern schedules: 1) a traditional "q3" call schedule with 30-hour recurrent shifts and 2) an intervention schedule that limited scheduled shifts to 16 hours but required four interns rather than the traditional three. Intern volunteers were randomized to work either the traditional schedule in the CCU and the intervention schedule in the MICU or the converse.

Results: Interns on the traditional schedule worked 19 hours more per week, slept 5.8 hours less per week, and had twice as many attentional failures while working at night ($p=0.02$). In addition, they made 36% more serious medical errors on the traditional schedule ($p<0.001$), including 21% more serious medication errors ($p=0.03$) and over five times as many serious diagnostic errors ($p<0.001$). In the units as a whole, there were 22% more serious medical errors on the traditional schedule ($p<0.001$).

Administrative Supplement: A medical simulator study of the effects of intern sleep deprivation on performance was conducted concurrently through an administrative supplement funded by AHRQ. Enrollment of subjects is complete, and results from this supplementary study are pending.

Key words: patient safety, medical errors, sleep deprivation, fatigue, schedule, intern, house staff, work hours, duty hours, performance

Purpose

The study was designed to assess the impact of work hours on intern sleep and patient safety in the intensive care unit (ICU) and Cardiac Care Unit (CCU) at a tertiary care teaching hospital. To understand the effects of intern sleep deprivation on patient safety, we conducted a comprehensive comparison of safety while interns worked on a traditional work schedule and an intervention work schedule that was designed to reduce sleep deprivation. Our goals were 1) to quantify work hours and sleep in interns during a traditional work schedule; 2) to compare subjective reports of work hours and sleep with simultaneous, independent, objective measures; 3) to measure the effect of an intervention designed to eliminate extended work shifts on physicians' work hours, sleep, and attentional failures; 4) to compare rates of serious errors on the two schedules in which interns were directly involved, as interns were the focus of our scheduling intervention; and 5) to compare overall rates of serious medical errors in order to evaluate the effects of intern schedules on the system as a whole.

Scope

Although data on medical errors in hospitals have been available for several years, the Institute of Medicine's report, "To Err Is Human," has galvanized efforts to improve the safety of the healthcare system in an unprecedented fashion. The report attracted the attention of Congress and the President; shortly after its release, the federal government issued a directive to implement safety improvements. The report found that errors in healthcare should be seen as the results of complex system failures rather than faults at the level of individual physicians, nurses, or others within the healthcare system. It argued persuasively that strong leadership and extensive systems redesign will be required to substantially decrease rates of errors. One area for potential redesign is the extended work schedules required of healthcare professionals, particularly junior house staff. The recurrent extended-duty on-call shifts of 30-38 hours in duration that recur every 3-4 nights induce both chronic and acute sleep deprivation, which are known to cause decrements in alertness and neurobehavioral performance and which may contribute to a higher risk of adverse events and medical errors and, consequently, reductions in patient safety.

Compared to other industries, medicine has a relatively high rate of serious errors and adverse events (AE). In 1984, the Harvard Medical Practice Study (MPS) found in a chart review of 30,000 admissions to New York State hospitals that 3.7% of hospitalized adult patients had an AE of medical therapy. The most common causes of adverse events were medication use, wound infections, operative complications, and diagnostic mistakes; 71% of AEs caused short-term disability, 3% caused permanently disabling injuries, and 14% caused death. Fully 69% of the iatrogenic injuries were judged to be preventable.

Despite the variable methodologies used to identify medical errors and adverse events associated with critically ill patients housed in intensive care units (ICUs), the error rate in these studies has been consistently high. The most comprehensive evaluation of ICU errors was an Israeli study that collected data by both direct observation and stimulated incident reporting. Donchin and colleagues defined human error as any deviation from standard conduct or actions related to standard operational instructions or routines of the unit; medical decisions were not included. Error reporting by physician and nursing staff was conducted by stimulated self-reporting; 24-hour continuous direct observation of patient activities was conducted on a randomly selected group of medical-surgical ICU patients. Activity was defined as any bedside patient interaction with the immediate surroundings; an average of 178 activities per patient per day was observed. The investigators found a mean of 1.7 errors per patient per day, including 46% committed by the physician staff, or 0.78 physician errors per patient day. Among the errors observed or reported, 29% were rated as severe or potentially life threatening.

In addition to being very harmful to patients, iatrogenic injuries are costly. The annual cost of AEs to New York State was estimated to be $878 million 1989 dollars in the Harvard Medical Practice Study. In a study of ADEs in a 700-bed hospital, adverse drug events (ADEs) and preventable ADEs were estimated to cost $5.6 million and $2.8 million per year, respectively. Extrapolating from prior studies, hospital-based ADEs are estimated to cost the nation $2 billion per year. Other studies have estimated the annual cost of drug-related morbidity and mortality in the US to be as high as $76.6 billion, with the majority, $47 billion, related to hospital admissions. This would slightly exceed the $45.2 billion spent on diabetes care each year.

Post-graduate physician education in the United States takes the form of a residency program in the individual’s chosen specialty and is almost completely hospital and patient-care oriented. In order to maximize exposure to patients and encourage continuity of patient care, residents are historically required to work extended work weeks and long overnight shifts. Prior to the recently implemented ACGME work-hour limitations, residents routinely worked up to 110 hours per week, with overnight shifts as long as 36-40 hours. Interns were required to stay in hospital one out of every three or four nights for many weeks or months. Even in New York, where resident work hours are limited by the Bell Commission regulations, 37% of all residents still report working more than 85 hours per week, and 20% of residents work more than 90 hours per week. It is likely that interns work some of the most extremely extended hours in our society.

Extended working hours of residents have been present for many years. Seventy years ago, internships were 2 years long, were unpaid, and required continuous residence in the hospital; marriage was prohibited. However, because treatment options were more limited and hospital patients were not, on average, as acutely ill as they are today, it was less common for resident physicians to remain awake all night.

Over the past 50 years, medical advances have greatly increased the workload of resident physicians, yet fundamental aspects of their work schedule remain unchanged, such as the 30- to 38-hour on-call shift every three or four nights.

There exists considerable institutional inertia when it comes to changing this training system. Some physicians who worked these extended schedules during their training do not want to see them changed, believing that long hours are appropriate for maximizing educational opportunities, continuity of care, and professionalism. Some have characterized internship and residency as a necessary right of passage into the elite healthcare profession, while others have even suggesting that it is an effective means of screening out those who are not well suited to the rigorous demands of medicine. Traditionalists argue that, in learning to be a competent and compassionate physician, interns need to be exposed to as many patients and diseases as possible and learn to provide continuity of care. They argue that working shorter hours will result in a loss of professionalism in physicians and that the bond between doctor and patient may be degraded.

Understanding the impact of this training program on the alertness and performance of physicians has previously been limited mainly due to methodological issues. Measuring surrogate outcomes of fatigue as they relate to patient care has proven problematic. Accordingly, efforts to link individual fatigue to adverse outcomes has been difficult. It is important to note, however, that there is no credible evidence to suggest that physicians are immune from the effects of sleep deprivation or that the patients under their care are not also vulnerable to fatigue-related errors. Similar safety lessons have been learned from other safety-conscious industries, most often after the occurrence of catastrophic events.

There are four physiological determinants of alertness and performance; circadian phase, number of hours awake (sleep homeostat), nightly sleep duration, and sleep inertia. Each of these four factors is adversely affected by the extended-duty schedules of interns and residents; consequently, performance is degraded. Thirty-hour on-call shifts result in misalignment of the circadian phase, cause acute sleep deprivation every 2 to 4 days, create chronic partial sleep deprivation that results in cumulative sleep debt, and promote situations in which performance on complex tasks is often required within minutes of awakening, when sleep inertia is highest. A systematic intervention to minimize physician fatigue and improve patient safety ideally addresses all four of these physiologic principles.

Methods

The Intern Sleep and Patient Safety Study was conducted as part of the Harvard Work Hours and Health Study from July 2002 to June 2004 in the medical intensive care unit (MICU) and coronary care unit (CCU) of a large academic hospital following approval by the Institutional Review Board. The MICU and CCU were selected for study because they are the rotations of this internal medicine training program with the longest work hours and because medical errors have been detected at higher rates in critical care than elsewhere. Both units have 10 adult critical care beds.

The study was conducted over 2 consecutive years and enrolled interns from two separate residency classes. Interns were informed about the study upon receiving their matching assignment in the spring of the fourth year of medical school. In year one, 24 interns were enrolled in the study, with 23 completing the entire protocol. In year two of the study, 24 interns were enrolled in the study, with 21 completing the entire protocol.

Design of Intervention Trial

After providing written informed consent, intern volunteers were randomized to work either the intervention schedule in the CCU and the traditional schedule in the MICU or the converse; these rotations were distributed throughout the year. The human research committee of Partners Healthcare and Brigham and Women's Hospital approved all procedures, and all participants provided written informed consent. Using a within-subjects design, we studied the interns during two 3-week rotations in the medical intensive care unit (MICU) and coronary care unit (CCU) while they followed a traditional schedule with extended work shifts of 30 consecutive hours every other shift or an intervention schedule in which work shifts were a maximum of 16 consecutive hours. During the traditional schedule, three interns provided continuous coverage on a 3-day schedule, consisting of a day shift (7 a.m. to 3 p.m.) on day 1 followed by an extended work shift from 7 a.m. on day 2 to 1 p.m. on day 3. The interns staffed weekly ambulatory clinics when they coincided with day 1 or day 3, and the average scheduled hours totaled 81 to 83 hours per week, depending on the clinic assignment. During the intervention rotation, four interns provided continuous coverage on a 4-day schedule, consisting of a standard day shift (approximately 7 a.m. to 3 p.m.) on day 1, "day call" on day 2 from 7 a.m. to 10 p.m. (the first half of the traditional extended work shift), and "night call" on days 3 through 4, from 9 p.m. on day 3 to 1 p.m. on day 4 (the second half of the traditional extended shift). The maximal scheduled duration of a shift was 16 hours. Interns only staffed clinics during day shifts (day 1); thus, the maximal number of scheduled work hours was 60 to 63 hours per week. To counter the effects of extended wakefulness before night work, interns were advised to take an afternoon nap before starting the night call. During the traditional schedule, no such opportunity was available owing to the requirement to work continuously during the day and night. Two weeks before each study rotation, the interns worked primarily on an ambulatory clinic.

During year one, 24 interns enrolled in the study, and 23 completed all aspects of the study; one withdrew during the second scheduled rotation. Three more participated in an intervention that ultimately was not used. Results from these four interns are not included in the within-subjects analyses.

During year two, 24 interns enrolled in the study, and 21 completed all aspects of the study. Three subjects withdrew for personal reasons.

Data Collection and Classification

Sleep Variables

All subjects were scheduled in an ambulatory clinic rotation prior to any ICU or CCU rotation in an effort to minimize sleep deprivation entering the study. A team of trained sleep technologists who provided continuous coverage during the study rotations accomplished sleep data collection. Interns completed a daily log recording details of sleep episodes. At least 3 days per week during MICU or CCU rotations, interns underwent continuous ambulatory polysomnographic (Vitaport-2/3, TEMEC Instruments) monitoring while at work or at home. During year one, on the basis of an average (±SD) of 334.5±33.4 hours of interpretable polysomnographic recordings with concomitant sleep logs per subject, 95.6±1.8 percent of the 30-second intervals, termed “epochs,” during which polysomnographic data were scored concurred with the sleep log entries for vigilance state. The total sleeping time per rotation derived from the two methods was also correlated across the 20 interns ($r=0.94$, $P<0.001$). The weekly duration of sleep was compared between the two schedules by within-subjects paired Student’s t-tests. The number of hours of sleep in the preceding 24 hours was calculated for each work hour and compared between rotation types by means of a chi-square test. Attentional failures were identified by means of continuous electrooculography (EOG) and defined as intrusion of slow-rolling eye movements into polysomnographically confirmed episodes of wakefulness during work hours. The number of slow eye movements recorded during all waking PSG epochs was determined by a single scorer in an unblinded fashion according to established criteria. All results were then validated in a blinded fashion by an independent scorer, who compared them with the rates recorded from 9 p.m. to 3 p.m. in a subgroup (10 percent) of EOG recordings. Throughout all testing periods, subjects wore actiwatches that allowed for the daily monitoring of activity and that have been validated to be a reliable measure of activity and sleep in field and laboratory studies. Furthermore, regular performance assessment was completed during work hours using the performance vigilance test (PVT). Mood assessments were completed using analogue scales. During on-call periods, hourly saliva was collected to measure salivary cortisol and melatonin in an effort to determine circadian phase of the subjects.

Patient Safety Variables

To measure patient safety under the two schedules, we developed an intensive data collection and evaluation methodology that expanded upon methodologies we have previously used in the study of medication errors. In this study, we focused on procedural and diagnostic errors in addition to medication errors. The definitions used to classify incidents are provided in the following table:

Medical Error	Any error in the delivery of medical care, whether harmful or trivial
Serious Medical Error	A medical error that causes harm or has significant potential to cause harm. This includes preventable adverse events, nonintercepted serious errors, and intercepted serious errors. Errors with little or no potential for harm are not serious errors, nor are nonpreventable adverse events.
Intercepted Serious Error	A serious medical error that is intercepted before reaching the patient
Nonintercepted Serious Error	A serious medical error that is not intercepted but does not cause detectable harm, despite reaching the patient
Adverse Event	Any injury due to medical management.
Nonpreventable adverse event	Unavoidable injury due to appropriate, error-free medical care
Preventable adverse event	Injury due to a nonintercepted serious error in medical management
Serious Medication Error	A serious medical error related to the ordering or administration of pharmaceuticals, blood products, or intravenous fluids
Serious Procedural Error	A serious medical error related to the performance of an invasive procedure, such as placement of a central venous or arterial catheter
Serious Diagnostic Error	A serious medical error related to the performance of a history or physical examination or to the ordering or interpretation of a diagnostic test

A team of two nurse chart reviewers and six physician observers collected data, supplemented by voluntary reports from clinical staff and a computerized events detection monitor. Direct observation was the principal means of detecting serious errors in which interns were directly involved; physician observers employed by the study followed study interns 24 hours/day, 7 days/week. All other data collection methods were designed to capture all serious medical errors – those in which interns both were and were not involved. Before beginning data collection, all staff received intensive training in the consistent, objective capture of data using standardized forms. Because it was not possible to blind data collectors to study condition, the importance of maintaining standardized procedures and remaining objective was reinforced repeatedly throughout the study.

Each suspected error or adverse event identified was independently rated by two physician investigators blinded to the identity of those involved and to whether the incident occurred on a traditional or intervention schedule. Reviewers categorized each incident as an adverse event, nonintercepted serious error, intercepted serious error, or error with little potential for harm. Reviewers rated the preventability of adverse events using a Likert scale (prevented, definitely preventable, probably preventable, probably not preventable, definitely not preventable); the preventability scale was collapsed to preventable/not preventable prior to analysis. Events deemed more likely due to underlying illness than medical therapy were excluded. Disagreements were resolved by discussion, following calculation of pre-discussion inter-rater reliability using the kappa statistic, as described below. The agreement among reviewers was very high: kappa = 0.90 for event categorization, and 0.80 for preventability.

Results

Results of the investigation are available for year one. Findings are presented in two categories: 1) the effects of the scheduling intervention on work hours and sleep and 2) the effects of the intervention on the occurrence of medical errors.

Work

Of the 20 interns who completed the study protocol, all worked longer during the traditional schedule (mean, 84.9±4.7 hours per week; range, 74.2 to 92.1) than during the intervention schedule (mean, 65.4±5.4 hours per week; range, 57.6 to 76.3; P<0.001). Seventeen of the 20 interns worked more than 80 hours per week during the traditional schedule, whereas all interns worked fewer than 80 hours per week during the intervention schedule. The average difference in work hours was 19.5 hours per week (range, 8.4 to 32.4), or 69.2 hours per rotation (range, 26.3 to 107.3). There was no correlation between an individual intern's work hours during the pre-ICU ambulatory clinic rotation and his or her subsequent ICU rotation (r=0.20, P=0.44 during the traditional schedule and r=0.20, P=0.43 during the intervention schedule) or even between an individual intern's ICU rotations (r=0.05, P=0.85). During the traditional rotation, over half of work shifts (133 of 223, or 60 percent) were extended (24 hours or more), and 85 percent of work hours (4,255 of 5,036) occurred during these shifts. Twenty-one percent of these work hours were logged after more than 24 hours of continuous duty. The intervention schedule had no extended work shifts, and 96 percent of work hours occurred within the 16 hours scheduled, in contrast to the traditional schedule, in which 58 percent of work hours occurred during extended shifts.

Sleep

Interns slept an average of 45.9±5.9 hours per week (6.6±0.8 hours per day) during the traditional schedule, 5.8 fewer hours per week than during the intervention schedule (mean, 51.7±6.0 hours of sleep per week, or 7.4±0.9 hours per day; P<0.001).

All but three interns slept more during the intervention schedule than during the traditional schedule. The weekly durations of sleep and work were significantly inversely correlated ($r=0.57$, $P<0.001$), with a predicted loss of 19.2 minutes of sleep per week for each additional hour of work per week. During the traditional schedule, 31 percent of work hours were preceded by 4 or fewer hours of sleep in the preceding 24 hours, and 19 percent of work hours were preceded by 2 or fewer hours of sleep in the previous 24 hours, compared with 13 percent and 6 percent, respectively, during the intervention schedule ($P<0.001$ for both comparisons). The percentage of work hours preceded by more than 8 hours of sleep in the prior 24 hours was 17 percent during the traditional schedule and was 33 percent during the intervention schedule ($P<0.001$). Interns reported taking a prophylactic nap before night call on 69.9±30.8 percent of occasions. On average, interns slept for 1.76±1.04 hours between 9 p.m. and 8 a.m. during the traditional schedule, significantly longer than they slept while working the corresponding hours during the intervention schedule (1.29±0.90 hours per shift, $P=0.02$).

The electrooculogram was monitored during waking hours for the presence of slow rolling eye movements, termed attentional failures, and assessed alertness while on duty. Attentional failures occurred at double the rate during the night (from 11 p.m. to 7 a.m.) on the traditional schedule than on the intervention schedule ($P=0.02$), and there was a trend toward an increased occurrence of attentional failures during the day as well (7 a.m. to 10 p.m.)

Medical Errors

In total, 2,203 patient days (traditional schedule: 1,294) were studied, representing 634 admissions to the units (traditional: 385; intervention: 249) and 5,888 hours of direct intern observation.

Patient and unit characteristics were very similar on the traditional and intervention schedules. Interns wrote similar numbers of orders per patient day and interpreted similar numbers of diagnostic tests. They performed 16.1% fewer procedures per patient day on the traditional schedule than on the intervention schedule. Length of stay and mortality did not differ significantly.

There were 35.9% more serious medical errors made by interns on the traditional schedule (136.0 per thousand patient days) than on the intervention schedule (100.1 per thousand patient days; $P<0.001$). Intercepted serious intern errors occurred 27.9% more frequently on the traditional schedule than on the intervention schedule (70.3 vs. 55.0 per thousand patient days, respectively; $P=0.02$). There were 56.7% more nonintercepted serious intern errors on the traditional schedule than on the intervention schedule (44.8 vs. 28.6 per thousand patient days, respectively; $P<0.001$). Preventable adverse events were not significantly different.

When broadening the assessment to include the entire care team and not just the intern team, the rate of all serious medical errors was 22.0% higher on the traditional schedule (193.2 per thousand patient days) than on the intervention schedule (158.4 per thousand patient days; P<0.001). Intercepted serious errors occurred 37.1% more frequently on the traditional schedule than on the intervention schedule (95.1 vs. 69.3 per thousand patient days, respectively; P<0.001). Rates of all nonintercepted serious errors did not differ significantly, nor did rates of preventable adverse events.

There was no significant difference in total adverse events (preventable plus nonpreventable) between the traditional and intervention schedules (85.0 and 93.5 per thousand patient days, respectively; P=0.31). Secondary analysis of serious medical errors in which interns were not involved revealed no significant differences (traditional, 40.2; intervention, 38.5 per thousand patient days; P=0.69).

Intern serious medication errors were 20.8% more frequent on the traditional schedule than on the intervention schedule (99.7 vs. 82.5 per thousand patient days, respectively; P=0.03). Interns' serious diagnostic error rates were more than five-fold greater on the traditional schedule than on the intervention schedule (18.6 vs. 3.3 per thousand patient days, respectively; P<0.001). Interns' serious procedural error rates were not significantly different.

Analysis of the types of all errors (intern errors plus errors in which interns were not involved) showed similar patterns. Serious medication errors occurred 17.1% more frequently on the traditional schedule than on the intervention (139.1 vs. 117.7 per thousand patient days, respectively; P=0.03). Serious procedural error rates did not differ significantly. Serious diagnostic errors were nearly twice as common on the traditional as on the intervention schedule (21.6 vs. 11.0 per thousand patient days, respectively; P<0.001).

Conclusion

This investigation provided the opportunity to accurately assess the impact of an intervention to reduce work hours. By incorporating physiologic principles that govern alertness, the intervention schedule reduced both total weekly work hours as well as the occurrence of extended (>24 hour) shifts. This reduction had a significant impact on the amount of total and nightly sleep obtained by the subjects. This rather basic finding counters a prevailing notion within medical training environments that less time at work does not translate into more sleep and better rested physicians. Furthermore, alertness, as measured by slow eye movements, was significantly impaired in subjects when working the traditional schedule compared to the intervention. This is consistent with the hypothesis that fewer extended work shifts lead to increased amounts of sleep and higher overall alertness.

Serious medical errors committed on the units were significantly lower during the intervention schedule compared to the traditional schedule, suggesting that an intervention that eliminates 30-hour-in-a-row shifts (a duration still allowed by the new ACGME duty-hour standards) could improve patient safety. When correlating scheduling changes, alertness, and medical errors, causality is difficult to establish. However, it was our *a priori* hypothesis that improvements in sleep resulting from elimination of extended-duration work shifts and reduction of work hours would lead to a decrease in serious medical errors. There were no differences in patient severity of illness or other individual or systemic variables that could independently account for the observed differences in medical error rates. Our randomized study design greatly diminished the likelihood of hidden confounding due to secular trend, seasonal effects, learning over the course of the year, or other external factors unrelated to our study. The sleep, alertness, and slow eye movement data provide a strong basis for the conclusion that attentional failures due to sleepiness played an important role in the rate or serious medical errors.

Limitations

The intervention schedule had limitations. Despite the fact that the extended work shift was split in half, most work shifts remained long enough to induce significant decrements in neurobehavioral performance owing to sleep deprivation and still exceeded the limits imposed by many other safety-sensitive industries. Moreover, the interns often had to rise between 4 a.m. and 6 a.m., the time of maximal sleep propensity and efficiency in this age group, to review their patients' progress before morning rounds. Because nearly a third of their work hours (31 percent) were preceded by 6 or fewer hours of sleep in the preceding 24 hours, they continued to carry a substantial sleep debt, accounting for the high residual rate of attentional failures on both schedules, even during the day. Furthermore, during both the traditional and the intervention schedule, reported work hours often exceeded both the scheduled weekly hours and the number of consecutive work hours scheduled, owing to the interns' obligation to ensure the continued care of their patients after their own shift was over. We feel that this is important when considering rational scheduling practices, as actual work hours often exceed scheduled hours. In addition, it is important to emphasize that not all interventions that reduce intern work hours will increase interns' sleep or improve patient safety.

Schedule design is a critical factor in determining the extent to which round-the-clock work schedules disrupt wake-sleep cycles, even when the number of weekly work hours remains the same. Furthermore, any systemic intervention that reduces work hours necessarily increases either provider workload (i.e., the number of patients covered by a provider at any time) or the number of hand-offs in care between medical personnel on shorter work shifts. Either can lead to increased error and adverse event rates. Night float systems, which use

residents on night shifts to allow physicians working extended-duration work shifts protected time for sleep, have their own set of risks. Night float residents often know patients less well than team members (particularly if multiple residents share responsibilities as night floats over the course of a week, or if night floats are responsible for an increased number of patients) and may themselves be sleep deprived and error prone. For these reasons, we ultimately decided not to implement a night float system as a means of reducing intern work hours, as originally planned. Our data support the hypothesis that elimination of extended-duration work shifts, in a system that minimizes cross-coverage, can improve patient safety. These gains might not be realized in systems that use extensive cross-coverage.

In examining error rates it is important to consider limitations as well. We studied two ICUs in a single hospital, and our results may not be generalizable to other settings. Although our study was very large compared to prior observational safety studies, the study was not powered to detect differences in preventable adverse event rates, which should be investigated in a larger scale, multicenter trial.

Another important limitation was our inability to blind the medical observers to the schedule of the interns, an issue commonly encountered in investigations of systemic patient safety interventions. We addressed this in two ways: first, we instructed observers – none of whom were study investigators – in the importance of consistent, objective detection of serious errors, regardless of study schedule. Second, all initial observations passed through a second review stage by two independent investigators who were blinded to study condition and who made final determinations of incident classification with extremely high reliability. Nonetheless, we cannot exclude the possibility that some bias may have resulted from the inability to blind the primary detection process, though our reliability data suggest that this bias was likely minimal.

Notably, our data on the high incidence of intercepted near misses in ICU settings indicate that the performance of personnel acting as an intern safety net—nurses, pharmacists, and senior medical staff—is very important in preventing actual injury secondary to intern errors. Therefore, future studies should seek to improve and measure objectively the sleep and performance of all clinical unit personnel, as team performance may critically affect safety. Having interns on a different schedule than supervising residents may have introduced discontinuities in education and interfered with the traditional resident-intern mentorship bond. We would recommend that future studies evaluate the effect of eliminating the extended-duration work shifts of both interns and senior residents, both to avoid this problem and because it is unlikely that interns are uniquely susceptible to the adverse effects of sleep deprivation.

Administrative Supplement

We receive an administrative supplement to fund a companion study investigating the effects of intern sleep deprivation on their performance of explicitly defined tasks on a medical simulator. We have completed full enrollment of subjects. Of 23 enrolled subjects, 17 completed all phases of the study, two subjects dropped out, and four partially completed the investigation. No adverse events occurred. A total of 61 observed sessions occurred over the 12-month period, during which 183 simulations occurred. Subjects were observed in a controlled fashion performing equivalent tasks in a rested, sleep-deprived, or a partially sleep-deprived state. Subjective feedback, as well as objective measures of medical errors and time to critical interventions, were measured by two investigators. All simulated sessions were videotaped.

Enrollment is now complete, and data analysis is ongoing. Blinded reviewer analysis of performance (on videotape) is ongoing, and data regarding medical errors are being accumulated. Intrasubject comparison will occur to document the effects of fatigue on medical decision making and performance. Intersubject comparison will also occur to investigate the types of errors most often made in rested or sleep-deprived state. We expect preliminary analysis to be complete in spring 2005.

Significance

This investigation provides much needed objective data regarding the impact of sleep deprivation on real-world performance in the healthcare training environment. Systematic improvements in scheduling that adhere to physiological principles and are aimed to increase sleep and improve performance may have implications for the health of critically ill patients. The prospective, randomized nature of the study provides a sound methodological basis to implement systematic changes in scheduling in order to improve patient safety.

To our knowledge, this is the first objectively validated data on work hours, sleep, and attentional failures in medical trainees *in situ* and quantifies the effects of eliminating extended-duration work shifts on these measures. These findings may apply not only to residents working in critical care units but also to those on other rotations and specialties and to more senior residents, attending physicians, nurses, and others.

Extended-duration work shifts and long work weeks continue to be permitted even under the scheduling reforms instituted last year by the Accreditation Council for Graduate Medical Education, and most of the 100,000 physicians who are currently in training in the United States work 30-hour shifts regularly. As such, extended work shifts may contribute to medical errors in teaching

hospitals. Future studies should evaluate the effects of current working practices on physicians and objectively measure the impact of interventions designed to reduce working hours on physician health, education, and safety, and patient safety.

Publications

Lockley SW, Cronin JW, Evans EE, Cade BE, Lee CL, Landrigan CP, Rothschild JM, Katz JT, Lilly CM, Stone PH, Aeschbach D, Czeisler CA, Effect of Reducing Interns' Weekly Work Hours on Sleep and Attentional Failures. *N Engl J Med* 2004;351:1829-37.

Landrigan CP, Rothschild JM, Cronin JW, Kaushal R, Burdick E, Katz JT, Lilly CM, Stone PH, Loclkey SW, Bates DW, Czeisler CA, Effect of Reducing Interns' Work Hours on Serious Medical Errors in Intensive Care Units. *N Engl J Med* 2004;351:1838-48.

Barger LK, Cade BE, Ayas N, Cronin JW, Rosner B, Speizer FE, Czeisler CA, Extended Work Shifts and the Risk of Motor Vehicle Crashes among Interns. In Press *N Engl J Med* 2005.