JCSM
Journal of Clinical Sleep Medicine

SCIENTIFIC INVESTIGATIONS

Working Memory Capacity is Decreased in Sleep-Deprived Internal Medicine Residents

Ashraf Gohar, M.D.[1]; Alexander Adams, M.P.H.[2]; Elie Gertner, M.D.[3]; Linda Sackett-Lundeen, M.T.(ASCP)[4]; Richard Heitz, Ph.D.[5]; Randall Engle, Ph.D.[6]; Erhard Haus, M.D.[5]; Jagdeep Bijwadia, M.D.[1]

[1]Division of Pulmonary, Allergy, and Critical Care, University of Minnesota and Regions Hospital, St. Paul, MN; [2]Section of Pulmonary and Critical Care, Regions Hospital, St. Paul, MN; [3]Section of Internal Medicine, Regions Hospital and Section of Rheumatology, University of Minnesota Medical School, Minneapolis, MN; [4]Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN; [5]School of Psychology, Vanderbilt University, Nashville, TN; [6]School of Psychology, Georgia Institute of Technology, Atlanta, GA

Background: Concerns about medical errors due to sleep deprivation during residency training led the Accreditation Council for Graduate Medical Education to mandate reductions in work schedules. Although call rotations with extended shifts continue, effects on resident sleep-wake times and working memory capacity (WMC) have not been investigated.

Objectives: The objective of this study was to measure effects of call rotations on sleep-wake times and WMC in internal medicine residents.

Methods: During 2 months of an internal medicine training program adhering to ACGME work-hour restrictions (between April 2006 and June 2007), residents completed daily WMC tests, wore actigraphy watches, and logged their sleep hours. This observational study was conducted during a call month requiring 30-hour call rotations every fourth night, whereas the noncall month, which allowed sleep/wake cycle freedom, was used as the control.

Main Outcome Measures: Sleep hours per night and WMC testing.

Results: Thirty-nine residents completing the study had less sleep per night during their call month (6.4 vs 7.3 h per night noncall, $p < 0.001$) and sleep per night varied from 3.7 to 10.1 hours. Call rotation caused greater self-assessed sleepiness and reduced WMC recall scores (-2.6/test, $p < 0.05$), and more math errors occurred when on call (+1.07/test, $p < 0.04$). Full recovery of WMC did not occur until the fourth day after call. On-call rotation on the first month had a confounding detrimental effect on WMC.

Conclusion: A month of call rotations reduced overall sleep per night; sleep hours per night were variable, and WMC was adversely affected. Decreased WMC could explain impaired judgment during sleep deprivation, although clinical error rates were not evaluated.

Keywords: Working memory capacity, sleep deprivation, residency training, shift work, wrist actigraphy

Citation: Gohar A; Adams A; Gertner E; Sackett-Lundeen L; Heitz R; Engle R; Haus E; Bijwadia J. Working memory capacity is decreased in sleep-deprived internal medicine residents. *J Clin Sleep Med 2009*;5(3):191-197.

A 1999 report from the Institute of Medicine estimates that approximately 100,000 people die in hospitals each year due to preventable medical errors—60% to 80% of the errors have been attributed to fatigue.[1] Responding to a probable association between errors and patient safety, and with expectations to improve resident health and create a better resident education experience, the Accreditation Council for Graduate Medical Education (ACGME) established requirements that limited resident working hours to fewer than 80 hours per week, restricted extended shifts to 30 hours or less, limited average call frequency to no more than every third night, and mandated a 10-hour rest period between shifts.[2] In 2008, the ACGME expanded the resident work-hour restrictions.[3] Recent studies have reported that limiting extended work shifts can decrease the risk of medical errors[4] and reduce the risk of postcall automobile crashes.[5] Neurobehavioral performance of residents after heavy call (90 hrs work/week + night call every 4th or 5th night) has been found to be equal to or worse than performing with a blood alcohol level of 0.04% to 0.05% during a light call (daytime clinic and backup night call) rotation.[6] Recent reports have found a variable effect on patient mortality rates after resident-hour reductions, ranging from no change to slight improvement.[7-9] But mortality rate is a blunt multifactorial endpoint that may not reflect ineffectiveness of an education experience or poor resident health status caused by sleep deprivation. Other studies conducted since the imposition of work-hour restrictions have reported improvements in rate of intensive care unit utilization, rates of discharge to home or rehabilitation facilities, number of pharmacist interventions to prevent errors,[10] and quality of care and efficiency in the coronary intensive care unit.[11] Nevertheless, the effect on resident performance after a 30-hour work shift has not been thoroughly studied.[12]

Submitted for publication October, 2008
Submitted in final revised form March, 2009
Accepted for publication March, 2009
Address correspondence to: Alexander Adams, Regions Hospital – Pulmonary Research, 640 Jackson ST, St. Paul, MN 55101; Tel: (651) 254-1390; Fax: (651) 254-3098; E-mail: alex.b.adams@healthpartners.com

Lack of sleep in residents has been found to affect mood, relationship-related stresses, and performance on simulated or standardized tests.[13] The mechanisms responsible for the effects of sleep deprivation on residents' performance have not been explored. Although prior studies have documented errors and adverse events associated with resident sleep deprivation,[4-5,13-14] this study explores changes in working memory capacity (WMC) in sleep-deprived residents as a potential mechanism leading to errors and adverse events.

WMC is the ability to retain and manipulate information or sensory input to perform multiple tasks[15] and is a combination of attention, concentration, and short-term memory.[16] Deficits in WMC have been shown to predict Alzheimer disease and the effects of alcohol consumption, as well as a range of cognitive phenomena.[17-18] Studies support the idea that sleep is important for memory consolidation[19] and that sleep deprivation results in decreased WMC.[20-21] A test for WMC has been standardized that is valid, reliable, and easy to administer.[22]

We hypothesized that sleep deprivation during a month of extended work shifts would cause a measurable decrease in WMC. Therefore, we measured WMC daily in internal medicine residents during call and noncall (control) rotations. Wrist-worn actigraphy, as well as sleep logs and self-reporting of sleepiness levels, were used to monitor sleep duration and its subjective effect. Specifically, we hypothesized that call rotations would reduce sleep duration—as measured by actigraphy—causing a reduction in WMC-recall scores and an increase in WMC-test errors.

METHODS

All residents in the internal medicine (PGY1-3) and medicine/pediatric (PGY 1-4) training programs at Regions Hospital, a teaching hospital affiliated with the University of Minnesota Medical School Training Program, were invited to participate in this study. The protocol was approved by the Institutional Review Board. To participate, all residents were required to sign an Institutional Review Board-approved informed consent. Each resident had to be scheduled within the year to call (hospital ward) and noncall (usually outpatient or certain subspecialty rotations) month-long rotations. Month order was randomized, as scheduling was set prior to study entry; therefore, start months were not necessarily consecutive or in a specific order, nor did we have control over scheduling. A 1-hour study-recruitment orientation session was presented to all residents at the beginning of their rotations at Regions Hospital. Participation was voluntary, and no residents were excluded if they were able to provide patient care during their call and noncall months. Upon completion of the study, residents earned the minor incentive of an iPod shuffle. The call month required up to 30-hour shifts every fourth night. The noncall month allowed residents to self-control their sleep-wake cycles. Both the call and noncall monthly rotations followed ACGME guidelines, that is, never exceeding 80 work hours per week.

Consenting residents completed a sleep questionnaire at the beginning of the study that included an Epworth Sleepiness Scale (ESS) assessment.[23] The residents completed daily standardized sleep logs and wore actigraphy watches continuously. Caffeine and alcohol intake were also recorded on the sleep log. The wrist actigraphs (MicroMini Motionlogger, Ambulatory Monitoring, Ardsley, NY) were collected weekly for data download, and, subsequently, the data were analyzed by Action 4 Software (Ambulatory Monitoring). Analysis involved eliminating times that the watch was taken off and determining a movement score per minute that exceeded a threshold for wakefulness. A calculation of sleep percentage per 24-hour period was determined from noon each day. Standardized WMC testing was performed daily between 11:00 and 14:00, as work allowed, in the resident's work area where orders and notes are written. This 20-minute test is the automated version of the operation span task (A-Ospan) test, which is a reliable computer-based test of WMC (Appendix).[22] Each testing session involved an instruction and practice period followed by taking the test in which residents attempt to solve 75 simple math problems in blocks of 3 to 7 problems per block. A response to each problem reveals a letter that must be recalled in the correct sequence for that block. WMC testing must have been completed on more than 50% of the in-house days for inclusion in data analysis. The residents also, subjectively, scored their sleepiness from 0 (alert) to 10 (very sleepy) before taking the daily WMC test. Morning-evening person tendency questionnaires were conducted[24] to assess possible associations with WMC testing results.

Based on data previously reported from a validation study[15] and using standard statistical assumptions for α and β, a minimum required sample size to achieve significant differences between months was estimated at 37 residents. The statistical analysis was conducted using the R-statistical programming language [25] employing a prospective, repeated-measures design in which each subject served as her or his own control. Five outcome variables were analyzed in separate multivariate models: Ospan score, letters recalled, math errors, accuracy errors, and speed errors. All of these scores range from 0 to 75 and were rescaled for analysis purposes so that higher-value scores implied better performance (errors were scaled to report number correct or number of errors). A separate analysis of the outcome variable, mean daily hours of sleep, was calculated from actigraph and sleep-log data. The key effects of interest were (1) call versus noncall rotation impact on the 5 outcome variables and (2) whether the call schedule was associated with a reduction in mean daily sleep time. An order effect of the call to noncall month compared to noncall to call month sequencing was also analyzed.

For the 5 Ospan outcomes, a separate multilevel regression model nesting Ospan scores within residents was fit for each outcome. The main effect of interest in the models was call versus noncall rotation. The models also were adjusted for 3 types of potentially confounding effects. First, they were adjusted for the possible confounding demographic effects of resident age and sex. Second, they accounted for mean daily caffeine intake, ESS score, and mean hours of sleep. Finally, the impact of rotation sequence (precall, on-call, postcall, post-postcall), time within rotation, prior rotation (on-call, no-call), and time between rotations was examined. Since residents missed testing on some days, we performed an analysis to ascertain if missing outcome variables represented a potential response bias—this was completed by a multivariate logistic regression model for each of the 5 outcomes.

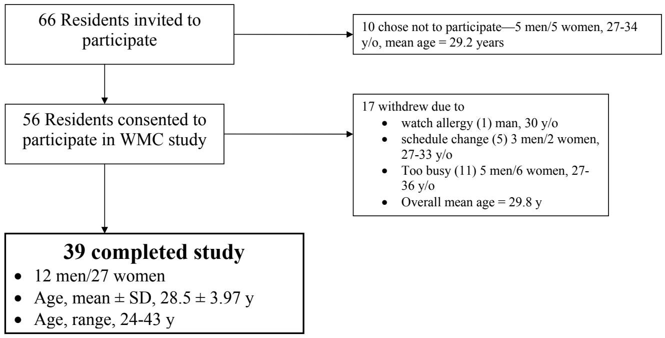

Figure 1—Residents who participated, refused, and withdrew from the study of working-memory capacity (WMC) and sleep. Of the 66 residents who were invited, 56 consented to participate during their call/noncall months. Seventeen residents withdrew due to a watch allergy, schedule changes, or becoming too busy to continue. Men tended to withdraw or refuse more than women. Overall, the 3 groups were similar in sex, age, and year of residency (55%-65% interns in each group).

The daily-hours-of-sleep difference between the call versus noncall months was analyzed by a linear regression model controlling for the confounding effects of age, sex, ESS score, daily caffeine intake, and starting noncall.

RESULTS

Between April 2006 and June 2007, 66 residents were invited to participate in this study. Of the 66 residents who were initially recruited, 39 completed the compliance requirements for WMC testing and sleep monitoring. Figure 1 displays characteristics of the refusals, those who withdrew with their reasons, and the participants. Differences between the 3 groups were not apparent other than a tendency for men to refuse or withdraw. The primary reason for withdrawing was the inability to complete WMC testing during the busy work schedule. Twenty-one started with the call rotation, and 18 started with the noncall rotation. The initial mean ESS score of the 39 participating residents was 9.8 ± 5.4. The Ospan test was completed by residents on 16.7 ± 3.66 days (70%) during the call rotation and 15.1 ± 2.73 (76%) of the days during the noncall rotation.

Table 1 displays the mean unadjusted Ospan scores, sleepiness scores, and sleep hours per day via actigraphy and sleep logs for residents during their call and noncall months. During the call-rotation month, the residents recalled fewer letters per test—their mean Ospan and letters scores were lower by 2.6, $p < 0.05$, and 1.7, $p < 0.05$. The number of math errors, speed errors, and accuracy errors were greater when residents were on call (Table 1, Figure 2). Mean reported visual-analog

Table 1—Results of Working-Memory Capacity Testing, Sleepiness Scores, and Hours of Sleep Per Day for the 39 Residents During their Call and Noncall Months

	Month		p Value
	Call	Noncall	
Ospan score	54.11 (1.91)	56.71 (1.54)	< 0.05
Correct letters	65.47 (1.07)	67.15 (0.75)	< 0.05
Math errors	5.32 (0.40)	4.25 (0.33)	< 0.04
Speed errors	1.24 (0.19)	0.76 (0.11)	< 0.03
Accuracy errors	4.07 (0.27)	3.49 (0.29)	< 0.03
Sleepiness scores	3.65 (0.14)	2.43 (0.16)	< 0.001
Sleep, h			
Actigraphy	6.39 (0.13)	7.25 (0.12)	< 0.001
Log	7.08 (0.11)	7.71 (0.09)	< 0.001

Data are presented as mean (SEM). The errors are reported as the mean number of errors per testing session.

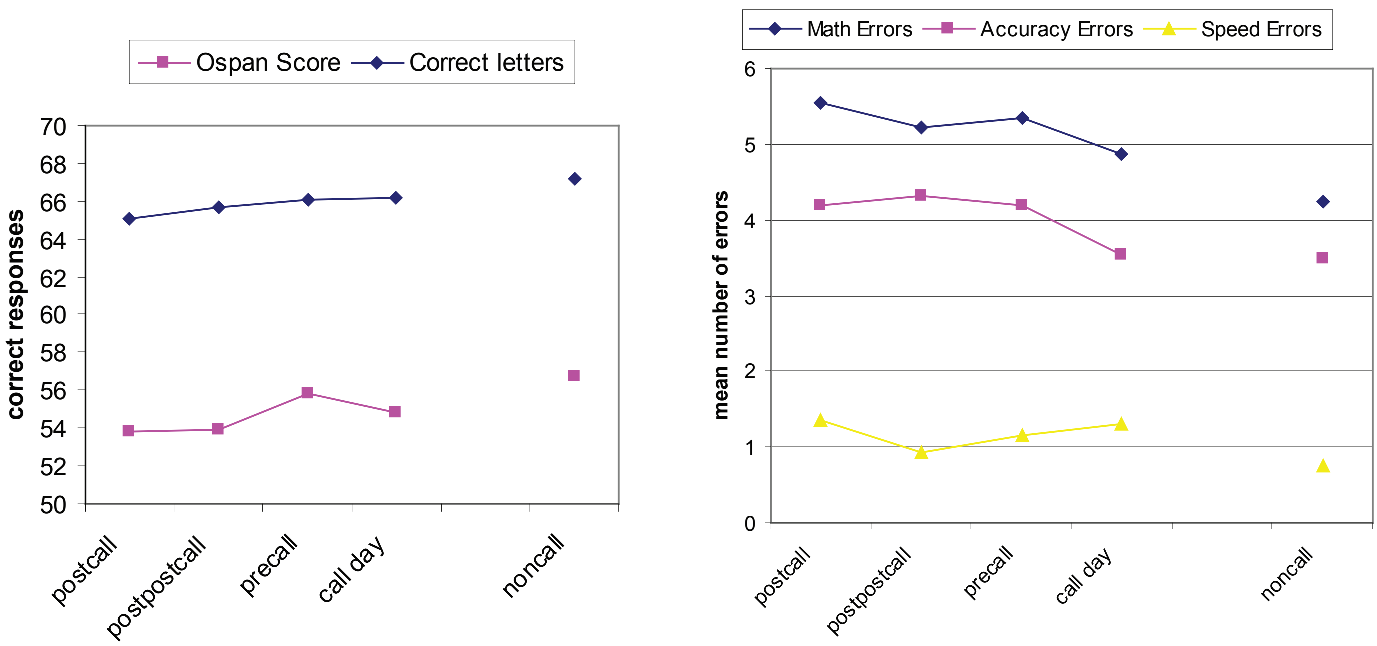

Figure 2—Mean daily working-memory capacity (WMC) test results (Ospan, Correct letters, Math errors, Accuracy errors, Speed errors) during the call-rotation cycle for immediate postcall, postpostcall, precall, and call days. Mean WMC scores were lower on the immediate postcall day, with gradual improvement until the call day. For comparison, mean daily WMC results for the entire noncall month are displayed on the right of each graph. Mean daily WMC scores during the call-month days never achieved the levels of the on-call month.

Table 2—Results of 5 Multivariate Models of WMC Tests

	Ospan Score		Letters		Math		Accuracy		Speed		
	Effect	p Value	Effect	p Value	Effect	p Value	Effect	p Value	Effect	p Value	Mean
Baseline WMC Score	66.845	< 0.001	76.119	< 0.001	74.640	< 0.001	76.356	< 0.001	72.220	< 0.001	
Male	7.001	0.041	3.665	< 0.001	0.508	0.45	0.381	0.44	0.086	0.78	31%
Age	-0.822	0.089	-0.708	< 0.001	-0.210	0.04	-0.227	0.00	0.048	0.30	28.13 y
Rotation	6.020[a]	< 0.001[a]	-0.009[a]	0.58[a]	2.399[a]	< 0.001[a]	1.252[a]	< 0.001[a]	1.122[a]	< 0.001[a]	
Sleepiness scoring at Time of Test	-1.360[a]	< 0.001[a]	0.171[a]	< 0.001[a]	-0.254[a]	< 0.001[a]	-0.124[a]	0.01[a]	-0.125[a]	< 0.001[a]	3.08
Day in rotation (1-31)	0.329[a]	< 0.001[a]	0.107[a]	0.76[a]	0.074[a]	< 0.001[a]	0.054[a]	< 0.001[a]	0.020[a]	< 0.001[a]	
Mean daily hours of sleep during rotation	0.242	0.089	2.705	0.10	-0.069	0.10	-0.077	0.01	0.010	0.71	7.20 h
Mean daily caffeine consumption	0.398	0.517	-0.544	0.03	-0.429	0.01	-0.276	0.03	-0.092	0.38	
ESS Score	0.141	0.617	0.118	0.40	0.038	0.52	0.065	0.13	-0.018	0.50	9.89
Began experiment on noncall rotation	10.294[a]	0.002[a]	4.842[a]	< 0.001[a]	2.762[a]	< 0.001[a]	1.486[a]	< 0.001[a]	1.336[a]	< 0.001[a]	
Noncall scores among those starting Noncall	-10.282[a]	< 0.001[a]	-5.531[a]	< 0.001[a]	-2.895[a]	< 0.001[a]	-1.331[a]	< 0.001[a]	-1.551[a]	< 0.001[a]	

WMC refers to working-memory capacity; ESS, Epworth Sleepiness Scale. [a]Significant.

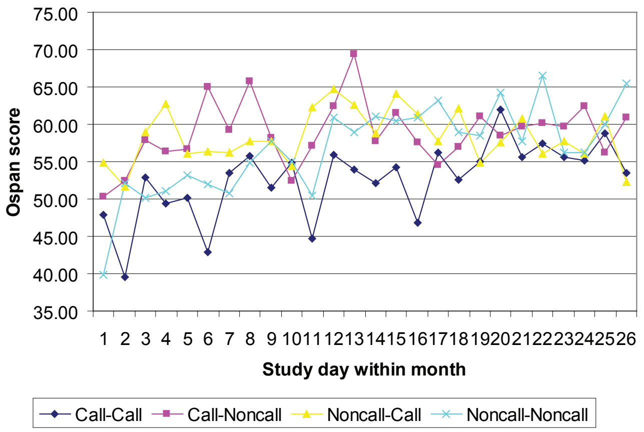

Figure 3—Temporal Ospan scores by call order. The mean Ospan scores are displayed for residents when assigned a call month first: during their call month (Call-Call, diamonds) and during their noncall month following the call month (call-noncall, squares). Scores are shown for residents assigned to their noncall month first**:** during their call month (noncall-call, triangles) following the noncall month and during their noncall month (noncall-noncall, xs). Note that a training or learning phase is apparent for each call-rotation group in which their scores gradually improve until approximately the 20th day of the rotation. An order effect was found to be significant—when the call month was assigned first, the Ospan scores for that month were lower than the other assignment combinations until late into the month.

sleepiness scores (0 = alert, 10 = very sleepy) were greater during the call rotation (3.65 ± 0.9 [mean ± SD] $p < 0.001$) compared with the non-call–rotation month (2.43 ± 0.99). The mean daily sleep duration during the call rotation (according to sleep log) was 7.08 hours, whereas sleep time was greater ($p < 0.001$), 7.71 hours, during the noncall rotation. Mean daily sleep duration by actigraphy was also significantly less during call compared with noncall rotations (6.39 hours vs 7.25 hours, $p < 0.001$).

The logistic-regression model showed no evidence that missing an Ospan score correlated with cognitive performance. For residents who missed testing sessions, missing values were not completely random, but there was no suggestion that a pattern of missingness was correlated with the outcomes of interest; therefore, analysis proceeded without adjustments for missing values.

The results of the final multivariate models of Ospan scores are contained in Table 2. Male sex and age were slightly associated with greater letter recall. Sleepiness and rotation significantly impacted WMC, as was seen in 2 variables: sleepiness scoring at time of test and rotation (on-call vs noncall). Sleepiness at time of testing was significantly associated with lower overall Ospan scores (-1.36, $p < 0.001$), and more Math errors (fewer correct) (-.254, $p < 0.001$), Accuracy errors (-.124, $p = .01$), and Speed errors (-.125, $p < 0.001$). In this statistical model, the mean daily hours of sleep during the rotation did not correlate with WMC; however, both mean daily hours of sleep and WMC scores were significantly decreased during the call compared with the noncall months (Table 1, t-test).

Learning occurred during the experiment, as is seen with increasing Ospan scores in Figure 3. Ospan scores increased a mean of 0.33 points ($p < 0.001$) per day during a given rotation. However, this learning differed by exposure to the test. Residents who began the experiment during an on-call rotation tended to score 10.29 points lower ($p = 0.002$) during both rotations after multivariate adjustment (Table 2). A similar effect was seen with the math, accuracy, and speed scores. The residents' rotation at the start of the study appears to have a lasting impact upon their Ospan performance. This effect can be seen in Table 3 and Figure 3. Residents starting the study during their call month had lower Ospan scores throughout the month, as well as intermittent decreases seen after their call nights (call-call, diamonds).

Table 4 displays the results of the multivariate model of the mean daily hours of sleep. Overall, residents slept 7 hours, 44 minutes per night. This did not appear to differ significantly by age or sex. It also did not appear to be impacted by ESS score, mean daily caffeine consumption, or starting on an on-call rotation. After adjusting for these factors, rotation (call vs noncall) was a significant factor on mean daily hours of sleep, with an effect of decreasing sleep by 35 minutes.

Table 3—Predicted Working Memory Capacity Scores

First Rotation-Current Rotation	Ospan	Letters	Math	Accuracy	Speed
Call-Call	56.74	61.40	61.40	65.69	71.71
Noncall-Call	67.04	64.17	64.17	67.17	73.04
Performance Difference	-10.29	-2.76	-2.76	-1.49	-1.34
Call-Noncall	62.76	63.80	63.80	66.94	72.83
Noncall-Noncall	62.78	63.67	63.67	67.09	72.61
Performance Difference	-0.01	0.13	0.13	-0.16	0.22

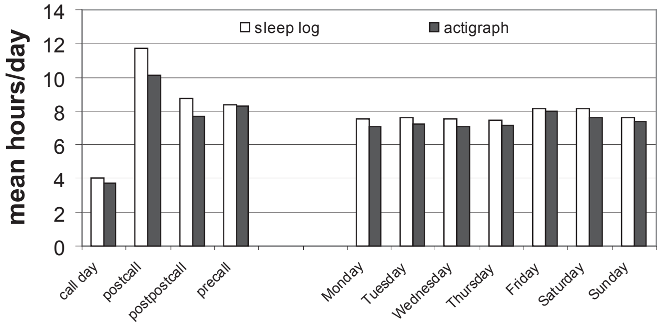

Figure 4—The mean variability in daily sleep time caused by call rotation, compared with noncall rotations. Daily mean sleep hours based on sleep logs and actigraphy during the 4-day call cycle rotation show decreased sleep on the call night followed by sleep payback during the subsequent 3 nights. The 4-day sleep-cycle pattern during the call rotation contrasts with the narrow range of sleep time per night during the noncall rotation. The noncall sleep pattern shows a comparatively minor payback on Friday and Saturday nights for, presumably, debt incurring during the week.

Table 4—Result of Multivariate Models of Mean Daily Hours of Sleep

	Effect	p Value
Mean daily hours of sleep	7 h 44 min	< 0.001
Age	-1 min	0.644
Male	-20 min	0.052
Rotation	35 min	< 0.001
ESS score	-1 min	0.984
Start on noncall rotation	-10 min	0.271
Mean daily caffeine consumption	-9 min	0.146

Multiple R-Squared: 0.272; Adjusted R-squared: 0.2068. ESS refers to Epworth Sleepiness Scale.

Ten residents scored slightly better on WMC testing during their call (vs noncall) rotation months. No clear factors, including age, ESS score, sleep duration, the number of caffeinated beverages, or "morning-evening person" tendencies could be identified that explain this ability. Although this finding is curious, only 3 residents recalled more than 5 letters or test better during their respective call month.

Actigraphy-based Sleep Cycles

Twenty-nine residents were compliant with actigraphy monitoring wearing, their watches 23.2 ± 3.3 days when on call, compared with 20.9 ± 4.3 days not on call. Data from 10 residents were excluded from the actigraphy monitoring because the residents were noncompliant with wearing the watch on an adequate number of days (> 50%). Actigraphy sleep hours were consistently less than sleep hours recorded by sleep logs (Figure 4). Sleep time included both naps and longer periods of consolidated sleep. The 4-day sleep-cycle pattern for the call-month rotation displayed in Figure 4 was distinctly different than the 7-day cycle for the noncall month. The mean number of actigraphy sleep hours for the "call" night was 3.7 hours, compared with 10.1 hours postcall. Call-month cycling was characterized by sleep reduction during the call day followed by payback for the following 3 days. The non-call–month pattern displayed little daily variation other than minor increases in amount of sleep per night on Friday and Saturday, presumably due to payback from debt accumulated during the week.

DISCUSSION

This was an observational study with residents serving as their own controls. We report the detrimental effect of a call rotation on WMC testing, as well as documenting more accuracy, speed, and math errors when residents are on a call schedule that requires extended shifts of up to 30 hours every fourth night. Associated measures of impaired clinical decision-making abilities or error rates were not evaluated during this investigation. Decreased sleep time during the call month was well documented by sleep logs and wrist-actigraph monitoring.

WMC may be a good marker of the potential mechanism for impaired judgment because the tasks of on-call residents are, by their very nature, complex; the residents are required to simultaneously assimilate, prioritize, and act on multiple sources of information in a stressful setting. This is particularly important because overnight and postcall errors in medication ordering have been reported to be higher, particularly in PGY1.[26] Eliminating interns' extended work shifts in an intensive care unit has been reported to significantly increase sleep and to decrease attention failures.[27] On the other hand, acute sleep deprivation in thoracic surgical residents has not been shown to affect operative outcomes.[28] Similarly 24 hours of call with acute fatigue does not worsen surgery residents' mood under the 80-work-hours-per-week regulations.[29] A paradoxical increase in sleep-latency time on the Multiple Sleep Latency Test has recently been reported in night-float residents.[30] Many studies have looked at the task performance of residents who are sleep deprived as the result of working extended shifts, but few studies have tested solutions other than eliminating extended shifts or allowing on-duty napping.[31]

Although the results of this study associate call rotation with WMC reduction in residents, an associated sleep deprivation itself may not be fully predictive of resident errors. We recognize that WMC deficit is only 1 of the possible mechanisms for causing errors; therefore, we can not speculate on the likelihood that call rotations will alter clinical decision-making abilities due to our reported effect on WMC.

This study provides solid and reliable sleep data collected over an extended period by sleep logs and verified by wrist actigraphy. The data document that overall sleep deprivation caused by call rotations affects WMC. This documentation of actual sleep deprivation is lacking from most studies of residents' sleep. Performance, automobile crashes, errors, mood, or stress have been evaluated after call,[4-5,13-14] but a measured overall lack of sleep, including the adverse effect of imposing a 4-day cycle of markedly altered sleep-wake times, has not been previously reported in studies of residents' sleep. Overall, the decreased mean number of hours of sleep during the call month did not associate with decreased WMC, but the addition of the highly variable sleep pattern accounted for in the "rotation" variable had a significant effect on WMC. To our knowledge, this is the only study showing that, when sleep debt incurred during the call night is not "paid off," residents remain in sleep debt during major portions of the rotation. We also observed a minor sleep-debt payoff during the non-call–month rotation, as seen by increased sleep time during weekends. This study provides examples of slow (noncall) vs rapid (call) sleep payoff, as has been previously described.[32]

ACGME rules have concentrated on minimizing sleep deprivation and subsequently improving safety by mandating an overall reduction in work hours. The detrimental effect of variable call schedules within the 80-hour–limit rule has not been rigorously evaluated.[12] Residents' schedules should be designed to provide safe care by clear-thinking residents, offer an optimal learning experience, and minimize detrimental health effects on the residents.

A subgroup of residents in this study did not show a decline in WMC testing during their call rotations. For these residents, WMC testing may not be sensitive to sleep-debt problems, or the residents may be resistant to the effects of call and are, therefore, better able to tolerate sleep deprivation. Previous studies have found variable effects of sleep deprivation.[33-34] Responses from our morning-eveningness questionnaire did not reveal an association with WMC results or other factors that might explain how these residents "beat the test." The potential mechanisms behind this phenomenon require further study. These residents may have similarities to the surgical residents who were reported to perform as well or better when tired.[28-29] The detrimental effects of sleep deprivation may, in fact, be dependent on an individual's vulnerability, as well as complex circadian and sleep-duration factors.

A number of limitations prevent a generalization of our results to other residency training programs or other settings. This was a single site study, limiting generalizability, and the overall compliance rate of this study was reduced at 70% to 75%. Unfortunately, the demands in the clinical areas did not allow residents to complete all requested WMC testing, but we speculate that testing when the residents are too busy or tired to attempt the tests would, in fact, have strengthened our findings. A significant confounding factor is the order effect identified in an interaction found between call rotation and starting the study during the call month. Residents starting the study on call had the burden of less sleep and a varying sleep-wake pattern while learning to perform the WMC testing, and we found a significant order effect (interaction) in our data analysis. Also, this study was not blinded—residents were aware of their rotation status during WMC testing, and that knowledge might have affected their interest in performing well on the tests, although that potential bias was not expressed. Although the results of this study are statistically definitive, we did not evaluate whether the deficiencies in WMC caused medical errors or otherwise impaired clinical decision making. Therefore, we cannot speculate on the importance of reduced WMC on clinical decisions, other than emphasizing this deficiency as a potential cause for impairment. To assess this association, an expanded study would be required that simultaneously evaluates WMC and clinical errors, such as incorrect prescription dosages.

In summary, residents experienced significant reductions in sleep hours during their call-month rotation. During call rotations, WMC recall rates were adversely affected, test errors increased, and self-reported sleepiness was greater. In addition, recovery from call required 3 nights of sleep to return to the baseline, and, on postcall nights, residents slept more than 10 hours before awakening. Our results support recent recommendations by the Institute of Medicine to extend rest periods after call and ensure at least 3 sleep nights between call rotations.

ACKNOWLEDGMENTS

This study was completed at Regions Hospital, St. Paul, Minnesota. To conduct this investigation, financial support was received from Healthpartners Research Foundation. No investigational or off-label medications were administered.

DISCLOSURE STATEMENT

This was not an industry supported study. The authors have indicated no financial conflicts of interest.

REFERENCES

1. Kohn L, Corrigan J, Donaldson M. To err is human: a safer health system. Washington, DC: Institute of Medicine; 1999.
2. Accreditation Council for Graduate Medical Education. Residents' duty hours. July 1, 2003. Accessed March 12, 2009, at http://www.acgme.org/DutyHours/dutyhoursummary2003-04.pdf.
3. Iglehart JK. Revisiting duty-hour limits—IOM recommendations for patient safety and resident education. N Engl J Med 2008;359:2633-5.
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;351:1838-48.
5. Barger LK, Cade BE, Ayas NT, et al. Extended work shifts and the risk of motor vehicle crashes among interns. N Engl J Med 2005;352:125-34.
6. Arnedt JT, Owens J, Crouch M, Stahl J, Carskadon M. Neurobehavioral performance of residents after heavy night call vs after alcohol ingestion. JAMA 2005;294:1025-33.
7. Shetty KD, Bhattacharya J. Changes in hospital mortality associated with residency work-hour regulations. Ann Intern Med 2007;147:73-80.
8. 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;298:984-92.
9. Volpp KG, Rosen AK, Rosenbaum PR, et al. Mortality among

hospitalized Medicare beneficiaries in the first 2 years following ACGME resident duty hour reform. JAMA 2007;298:975-83.

10. Horwitz LI, Kosiborod M, Lin Z, Krumholz HM. Changes in outcomes in internal medicine inpatients after work-hour regulations. Ann Intern Med 2007;147:97-103.
11. Bhavsar J, Montgomry D, Li J, et al. Impact of Duty hours restrictions on quality of care and clinical outcomes. Am J Med 2007;120:968-74.
12. Ciolli A. The medical resident working hours debate: a proposal for decentralized regulation of graduate medical education. Yale J Health Policy Law Ethics 2007;7:175-228.
13. Fletcher KE, Underwood W, Davis SQ, Mangrulkar RS, McMahon LF, Saint S. Effects of work hour reduction on residents' lives. JAMA 2005;294:1088-100.
14. Cappuccio FP, Bakewell A, Taggart FM, et al. Implementing a 48-hour EWDT-compliant rota (rotation) for junior doctors in the UK does not compromise patient's safety: assessor-blind pilot comparison. QJM 2009:102:271-82.
15. Heitz RP, Unsworth N, Engle RW. Working memory capacity, attention control and fluid intelligence. In: Wilhelm O, Engle RW, eds. Handbook of Understanding and Measuring Intelligence. London: Sage Publications Inc., 2004:61-78.
16. Budson AE, Price BH. Memory dysfunction. N Engl J Med 2005;352:692-9.
17. Calderon J, Perry RJ, Erzinclioglu SW, Berrios GE, Dening TR, Hodges JR. Perception, attention, and working memory are disproportionately impaired in dementia with Lewy bodies compared with Alzheimer's disease. J Neurol Neurosurg Psychiatry 2001;70:157-64.
18. Gotham AM, Brown RG, Marsden CD. Frontal cognitive functions in patients with Parkinson's disease on and off levodopa. Brain 1988;111:299-321.
19. Papgno C, Valentine T, Baddeley A. Phonological short-term memory and foreign-language vocabulary learning. J Mem Lang 1991;30:331-47.
20. Lee HJ, Kim L, Suh KY. Cognitive deterioration and changes of P300 during total sleep deprivation. Psychiatry Clin Neurosci 2003;57:490-6.
21. Smith ME, McEvoy LK, Gevins A. The impact of moderate sleep loss on neurophysiologic signal during working-memory task performance. Sleep 2002;25:56-66.
22. Unsworth N, Heitz RP, Schrock JC, Engle RW. An automated version of the operation span task. Behav Res Methods 2005;37:498-505.
23. Johns MW. Sleepiness in different situations measured by the Epworth sleepiness scale. Sleep 1994;17:703-10.
24. Horne JA, Ostberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol 1976;4:97-110.
25. R Development Core Team (2007). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Accessed March 12, 2009, at http://www.R-project.org.
26. Hendey GW, Barth BE, Soliz T. Overnight and post-call errors in medication orders. Acad Emerg Med 2005;12:62934.
27. Lockley SW, Cronin JW, Evans EE, et al. Effect of reducing interns' weekly work hours on sleep and attentional failures. N Engl J Med 2004;351:1829-37.
28. Ellman PI, Kron IL, Alvis JS, et al. Acute sleep deprivation in the thoracic surgical resident does not affect operative outcomes. Ann Thorac Surg 2005;80:60-4.
29. Kiernan M, Civetta J, Bartus C, Walsh S. 24 hours on-call and acute fatigue no longer worsen resident mood under the 80-hour work week regulations. Curr Surg 2006;63:237-41.
30. Surani S, Subramanian S, Aguillar R, et al. Sleepiness in medical residents: impact of mandated reduction in work hours. Sleep Med 2007;8:90-3.
31. Arora V, Dunphy C, Chang VY, Ahmad F, Humphrey HJ, Meltzer D. The effect of on-duty napping on intern sleep time and fatigue. Ann Intern Med 2006;144:792-8.
32. Drake CL, Roehrs TA, Burduvali E, et al. Effects of rapid vs slow accumulation of eight hours of sleep loss. Psychophysiology 2001;38:979-87.
33. Durmer JS, Dinges DF. Neurocognitive consequences of sleep deprivation. Semin Neurol 2005;25:117-29.
34. Mu Q, Mishory A, Johnson, et al. Decreased brain activation during a working memory task at rested baseline is associated with vulnerability to sleep deprivation. Sleep 2005;28:433-46.

APPENDIX

Automated Ospan Test

The automated Ospan (A-Ospan) is a computerized test that includes items (letters) to remember and a distracting activity in the form of problem solving. The A-Ospan consists of a basal part and the actual test. The basal part includes 12 problems to solve without letters to memorize, and a mean time used to solve the 12 equations is calculated to determine when equations will disappear during the actual test (for establishing speed error criteria). The actual test consists of 75 simple math problems with 75 letters to recall. The problems and letters are blocked, with each block containing a number of problems (between 3 and 7), with 1 letter to recall following each problem. At the end of each block, a screen appears, and the resident attempts to select the letters in the sequence in which they were revealed following the math problems. The number of correctly recalled letters determines the letter score, and recalling letters in the correct sequence determines the Ospan score. Incorrectly solved problems count as accuracy errors, and problems not solved in time are speed errors. Math errors are the sum of the accuracy and speed errors. The perfect score is Ospan score 75, total correct letters 75, math errors 0, speed errors 0, and accuracy errors 0.