Sleep Loss and Fatigue in Residency Training
A Reappraisal

Sigrid Veasey, MD
Raymond Rosen, PhD
Barbara Barzansky, PhD
Ilene Rosen, MD
Judith Owens, MD, MPH

LONG WORK HOURS ARE A TIME-HONORED TRADITION IN MOST RESIDENCY PROGRAMS. Demanding schedules are often said to be necessary for learning and development of professionalism. The use of resident physicians to provide relatively inexpensive coverage has also become an important economical factor for teaching hospitals within the United States. Recent studies of the negative effects of sleep loss and fatigue, however, call this model into question.

Any consideration of the impact of sleep loss and fatigue on medical training and patient care should be based on relevant data. In this article, we concentrate on 3 questions. First, is the quality of patient care jeopardized when residents are deprived of sleep? Second, is the well-being of residents at risk when they have not had adequate sleep? Third, what strategies are effective for overcoming the negative effects of sleep loss? We answer these questions by summarizing data on the effects of sleep loss and fatigue on cognitive and performance skills of residents. Reduced sleep time is commonplace for many interns and residents. Recent studies, however, suggest that sleep loss and fatigue result in significant neurobehavioral impairments in healthy young adults. We reviewed studies addressing the effects of sleep loss on cognition, performance, and health in surgical and nonsurgical residents. We describe the effectiveness of countermeasures for sleepiness, including recent work-hour restrictions. A more complete understanding of the issues of sleep loss during residency training can inform innovative strategies to minimize the effects of sleepiness and fatigue on patient care and resident safety.
tasks) and serial mathematical calculations were equally affected by 24 hours of total sleep loss and 1 week of sleep restriction to 5 hours of sleep per night.13-14 A meta-analysis of laboratory-based sleep loss studies estimated that the mean cognitive performances of healthy young adults who are sleep deprived (both short-term and chronic) are 1.3 SDs or more below the mean.8

Contrary to a popular belief that healthy adults can acclimate to sleep loss, the effects of chronic partial sleep loss appear to be cumulative.9-11 Specifically, sleepiness has been found to increase8 and performance on tests of vigilance and mathematical calculations to decline across 7 days of 5 and even 7 hours of sleep per night.10,11 Subjects often underestimate their own degree of sleep-related impairments in vigilance after 1 week of partial sleep restriction.9,12 Thus, they may mistakenly believe that they have acclimated to sleep deprivation.

In addition to reduced vigilance, verbal processing and complex problem solving13,15 are impaired with both short-term and chronic partial sleep loss. Learning for both complex cognitive and procedural tasks can decrease by up to 50% when sleep loss occurs between trials.14-16 Although health care professionals are sometimes assumed to be less susceptible to these effects, this has yet to be demonstrated.

Other variables may also affect performance after sleep loss, adding to impairments related to homeostatic influences.9 For instance, serial addition testing is poorest and lapses in vigilance after sleep loss are most frequent at the circadian nadir, between 2 AM and 9 AM.5 After sleep loss, tasks of longer duration are more likely to reveal sleep loss decrements in performance than are tests of shorter duration. Although subjects may not demonstrate impairment on brief tests (<10 minutes) following 24 hours of sleep deprivation,7-10 their performance may be impaired if the test duration is extended to more than 30 minutes.11 Caffeine, other stimulants, and physical exertion promote alertness and can improve measures of vigilance.12-24 Warm ambient temperature,25 reduced body temperature,26 and recent food intake27 can all increase sleepiness. Thus, studies evaluating the effects of sleep loss on neurobehavioral performance must be carefully controlled for such factors.

EFFECTS OF SLEEP LOSS ON RESIDENT PERFORMANCE

We conducted a MEDLINE search for peer-reviewed studies on the effects of sleep loss on resident performance measures for the years 1966 to 2002, using the search terms on-call, sleep, sleep loss, sleep deprivation, fatigue, intern, resident, surgeon, surgical, housestaff, and physicians in training. This search revealed a total of 50 articles. There were 33 studies reporting measured effects of sleep loss on physician cognitive or professional performance.58-60 In addition, there were 4 review articles,61-64 13 surveys or sleep log studies,65-77 1 electroencephalography study,78 and 1 study analyzing dexamethasone suppression tests and psychological profiles in residents.79 The table summarizes results from the 33 performance studies and provides a characterization of study design and sample sizes along with the quality of the studies and the incorporation of appropriate controls.

Surgical Residents

The MEDLINE search identified 10 studies evaluating surgical resident performances before and after either short-term sleep loss or on-call sleep loss.20-27 None of the studies included a control group of subjects with unrestricted or normal sleep for at least 1 week prior to study. This is important because surgical residents average less than 6 hours of sleep per night, even when overnight call shifts are scheduled every fourth night.60 Thus, the 10 studies have been designed to identify differences in performance attributable to acute sleep deprivation (usually defined as <4 hours of sleep for the night prior to testing) superimposed on chronic partial sleep loss of 5 to 6 hours per night.

Overall, the studies in surgical house staff show few additive effects of short-term sleep loss on surgical residents' cognitive performances. Reading comprehension and retention of clinical information was not affected in a crossover trial of “rested” and nonrested conditions in 46 surgical residents.33 However, the amount of sleep in the 2 nights prior to a test of longer duration, the American Board of Surgery In-Training Examination (ABSITE), was found to contribute approximately 7% of the variance in ABSITE scores. Deaconson et al30 studied 26 surgical residents under similar rested and unrested conditions, controlling for time of day, learning effect, and motivation. Residents' performance on a 1-hour battery of psychomotor tests was unaffected by acute sleep loss. (One of the tests in which surgeons did not show a deterioration with acute sleep loss, the Trail-Making Test,20 did identify significance performance decrements in medical house staff after sleep loss.56 The Trail-Making Test is a vigilance test in which subjects must connect dots using an alternating pattern of numerical and alphabetical order.)

Another study of 42 surgical residents found increased subjective feelings of anger, confusion, and fatigue after sleep loss, although there were no differences on cognitive measures.31,32 Reznick and Fosse39 validated several neurobehavioral assays for testing surgical residents' performances and also found minimal effect of acute sleep loss. Several of these studies had significant numbers of dropouts and residents who declined the study (see the Table). The reasons for dropout are generally not provided; it is possible that these reasons were confounded with vulnerability to sleep loss.

The manual dexterity and surgical skills of surgical residents may be specifically vulnerable to the effects of short-term sleep loss. Goldman et al20 compared videotapes of operations performed by surgical residents after less than 2 hours of sleep with videotapes of the same residents performing operations after longer sleep bouts. As the
Table. Summary of Studies Included in Review*  

<table>
<thead>
<tr>
<th>Source</th>
<th>Study Design</th>
<th>Objective Measures (Test Sensitive to Sleep Loss?)</th>
<th>Findings</th>
<th>Statistical Power for Null Shown</th>
<th>Circadian Time</th>
<th>Chronic Sleep Debt</th>
<th>Sleep Quantified</th>
<th>Stimulant, Alcohol, or Medications Excluded</th>
<th>Adequate Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goldman et al,28 1972</td>
<td>Observational (n = 5)</td>
<td>Qualitative operative technique (Y)</td>
<td>Poorly planned maneuvers 30% of operating time</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Reznick and Folse, 29 1987</td>
<td>Randomized testing order; paired t tests (n = 33)</td>
<td>40-item factual recall test (N); Identify abnormal laboratory values in a 48-page document (N); Suture placement in cloth (N)</td>
<td>No differences observed for any of the tests</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Deaconson et al,30 1988</td>
<td>Serial repeated measures (n = 26)</td>
<td>PASAT Addition (Y); Trail-Making (Y); Grammatical Reasoning (Y); Spatial Problems (N); Purdue Pegboard (N)</td>
<td>Statistically significant reduction in PASAT and Grammatical Reasoning scores</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Bartle et al,31 1988</td>
<td>Paired residents, crossover (n = 4 after 8 dropouts)</td>
<td>Digt symbol (Y); Story recall (N); Digt vigilance (Y); PASAT (Y); Raven’s Matrices (Y); Delayed recall (N); Pegboard (N)</td>
<td>No differences observed for any of the tests</td>
<td>N</td>
<td>Not AM</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Light et al,32 1989</td>
<td>Same subjects and design as in Bartle et al31</td>
<td>Pegboard (N)</td>
<td>PGY-1 performance decline with sleep loss</td>
<td>N</td>
<td>Not AM</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Browne et al,33 1994</td>
<td>Multiple comparisons (n = 4 after 22 incompletes or dropouts)</td>
<td>Short- and long-term recall of medical literature (N)</td>
<td>No difference</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Haynes et al,34 1995</td>
<td>Retrospective study of &gt;6000 surgical cases</td>
<td>Frequency of perioperative complications (N)</td>
<td>For operations on noncall days, 45% more likely to have complications if surgical resident is postcall (P&lt;.002)</td>
<td>Y</td>
<td>NA</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>?</td>
</tr>
<tr>
<td>Taffinder et al,35 1998</td>
<td>Prospective within-resident design (n = 6)</td>
<td>Errors in procedural skill (Y); Time to complete task (Y)</td>
<td>20% more errors (P&lt;.01); 14% more time (P&lt;.001)</td>
<td>?</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Grantcharov et al,36 2001</td>
<td>Prospective (n = 14)</td>
<td>Errors in procedural skill (Y); Time to complete task (Y)</td>
<td>2-fold increase in errors (P = .006); 38% increase in time (P = .002)</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Codellias and Huang,37 2001</td>
<td>All 34 surgical residents in 1 residency program who had taken the ABSITE twice</td>
<td>Questionnaire and ABSITE score (Y)</td>
<td>Regression analysis showed that the amount of sleep in the 2 nights prior to testing explained 7% of the test score</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Friedman et al,38 1971</td>
<td>Prospective paired t test (n = 14)</td>
<td>Identification of dysrhythmias on ECG strip for 20 minutes (N)</td>
<td>9 errors after sleep loss vs 5 after rested night (P&lt;.01)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Poulton et al,39 1978</td>
<td>30 junior hospital physicians</td>
<td>Grammatical Reasoning (Y); Reviewing laboratory sheets</td>
<td>Significantly less efficient with 8 hours of sleep loss</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Christensen et al,40 1977</td>
<td>Radiology residents, paired t test (n = 14)</td>
<td>Identification of chest radiograph nodule (N)</td>
<td>No difference; 61% of nodules identified in fatigued and rested states</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>

* (continued)
## Table. Summary of Studies Included in Review* (cont)

<table>
<thead>
<tr>
<th>Source</th>
<th>Study Design</th>
<th>Objective Measures (Test Sensitive to Sleep Loss?)</th>
<th>Findings</th>
<th>Confounders Controlled For</th>
<th>Statistical Power for Null Shown</th>
<th>Circadian Time</th>
<th>Chronic Sleep Debt</th>
<th>Sleep Quantified</th>
<th>Stimulant, Alcohol, or Medications Excluded</th>
<th>Adequate Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bertram,41 1988</td>
<td>10 second-year residents, randomly sampled; multivariate regression of patient encounter with history of on-call schedule</td>
<td>Number of items on history, physical examination, and tests ordered (N); clinical acumen not tested</td>
<td>No significant relationship between call schedule and number of items on history, physical examination, or number of tests ordered</td>
<td>?</td>
<td>?</td>
<td>N</td>
<td>N</td>
<td>?</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Ford and Wientz,42 1984</td>
<td>Psychiatry and medicine volunteers (31% of interns)</td>
<td>Reaction time and critical Flicker Fusion (Y)</td>
<td>Reaction time performance decreased with decreased sleep time in past 7 days (P&lt;.05); Flicker Fusion test performance improved with less sleep (P&lt;.05)</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Klose et al,43 1985</td>
<td>14 senior anesthesia residents tested for 5 consecutive days with varying on-call schedules (within-subject design)</td>
<td>Card Sort (Y) Digit Symbol (Y) Color Word (Y) Pegboard (Y)</td>
<td>No effect from on-call scheduling</td>
<td>?</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Denisco et al,44 1987</td>
<td>21 anesthesia residents; within-resident randomized design; rested vs post-call</td>
<td>Two 30-minute videotapes simulated anesthesia with “distractor” videotaped lecture (N)</td>
<td>Significantly lower vigilance scores in fatigued state</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Engel et al,45 1987</td>
<td>Crossover (n = 7)</td>
<td>Simulated new patient clinic visits; history and write-up (N)</td>
<td>No significant differences in history taking or write-up</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Orton and Gruzelier,46 1989</td>
<td>Paired t test; junior physicians (n = 20)</td>
<td>Reaction Time (Y) Haptic Sorting (N)</td>
<td>Significant deteriorations in both measures by 25% (P&lt;.001)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Storer et al,47 1989</td>
<td>Pediatric residents (n = 45), rest after sleep loss</td>
<td>30-question board test (N) IV placement (cats) (N) A-line (cats) (N) Intubation (cats) (N)</td>
<td>Test scores nonsignificant; IV, nonsignificant; A-line, 4 times longer to place (P&lt;.05)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Robbins and Gottlieb,48 1990</td>
<td>23 medical house staff</td>
<td>Simulated driving (Y) Rapid # Comparison (N) Spatial perception (N) Simple math (N)</td>
<td>Significant impairments in all but math</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Jacques et al,49 1990</td>
<td>Retrospective; family practice residents (n = 353)</td>
<td>American Board of Family Practice Examination, complete examination</td>
<td>Statistically significant negative correlation between score and last night’s sleep</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Ruben et al,50 1991</td>
<td>Medical residents (n = 63); paired t test; 36-hour call</td>
<td>Continuous performance (Y) Hand-eye coordination (N) Digit Symbol (Y) Horizontal addition (N)</td>
<td>All deteriorated significantly (P&lt;.001), except hand-eye coordination, which improved</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Nelson et al,51 1995</td>
<td>Anesthesia residents (n = 9); 20- vs 520-minute sleep</td>
<td>Torrance Test of Creative Thinking (N)</td>
<td>Less (P&lt;.05) verbal fluency, originality, and flexibility after sleep loss</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
### Table. Summary of Studies Included in Review (cont)

<table>
<thead>
<tr>
<th>Source</th>
<th>Study Design</th>
<th>Objective Measures (Test Sensitive to Sleep Loss?)</th>
<th>Findings</th>
<th>Statistical Power for Null Shown</th>
<th>Chronic Sleep Debt</th>
<th>Nap Quantified</th>
<th>Stimulant, Alcohol, or Medications Excluded</th>
<th>Adequate Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicine, Family Practice, Pediatrics, and Residents Radiology (cont)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hart et al.(^{22}) 1987</td>
<td>30 of 37 medical interns randomly assigned to postcall vs night at home</td>
<td>Sternberg short-term memory (N) Auditory serial addition task (Y) Delayed recall stories (N)</td>
<td>Overall significant differences (P&lt;.001); longer response latencies (10%); impaired story retention 77% vs 90%</td>
<td>?</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Gottlieb et al.(^{13}) 1991</td>
<td>32 PGY-1 and PGY-2 internal medicine residents, no dropouts; medical skills compared for long call and night float call schedules</td>
<td>Chart review study and medical errors detected by pharmacy (N)</td>
<td>Increased total sleep time on call (P&lt;.001); patients' length of stay, shorter with night float (P&lt;.01); fewer laboratory studies/radiographs ordered; 25% fewer medication errors (P&lt;.02)</td>
<td>?</td>
<td>?</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Lingen Felser et al.(^{14}) 1994</td>
<td>Prospective study of 40 junior physicians paired for sleep deprivation and controlled sleep</td>
<td>Number connection test, Stroop, reaction time, ECG accuracy, mood profile (Y)</td>
<td>All tests showed significant impairments following sleep loss</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Hawkins et al.(^{23}) 1995</td>
<td>68 medicine residents: 30 classified sleep-deprived (acute, +/- chronic), 38 rested; comparison self-report behavioral variables, psychomotor and cognitive task</td>
<td>20-minute battery Raven's Matrices (Y) Trail-Making (Y)</td>
<td>Definition “sleep-deprived” adjusted from &lt;4 hours to &lt;5 hours after original analysis nonsignificant; main effects on performance acute but not chronic sleep deprivation</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Self-report</td>
<td>Asked</td>
</tr>
<tr>
<td>Leonard et al.(^{24}) 1998</td>
<td>16 preregistration medical house officers randomly recruited; precall vs completing 32-hour shift</td>
<td>Delayed story recall (N) Critical Flicker Fusion test (Y) Trail-Making Stroop Coder-Word (Y) Grammatical Reasoning (N)</td>
<td>Significant differences in Stroop and in Trail-Making (P&lt;.05); no difference in story recall, Flicker Fusion, or Grammatical Reasoning</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>?</td>
</tr>
<tr>
<td>Richardson et al.(^{25}) 1996</td>
<td>26 interns recruited from unspecified number of medical interns; 2-group analysis of variance night float covered vs on call</td>
<td>Divided Attention Task (Y)</td>
<td>No difference in total sleep times for the 2 groups; no difference in performance</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Smith-Coggins et al.(^{26}) 1994</td>
<td>6 ED attendings tested during day and night shift</td>
<td>Mood ratings, serial performance tests: simulated triage task (Y) and simulated intubation (Y)</td>
<td>Night shift associated with intubation inefficiency and deterioration in performance over shift, no difference in accuracy</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Smith-Coggins et al.(^{26}) 1997</td>
<td>6 attending ED physicians; repeated measures analysis for night-shift vs day-shift performance</td>
<td>Vigilance reaction time (Y) Intubation of a mannequin (N) ECG analysis (N)</td>
<td>Night-shift sleep times were 1 hour less (P&lt;.05); reaction time and time to intubate were significantly slower for night shift (P&lt;.05)</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Self-report</td>
<td>N</td>
</tr>
<tr>
<td>Deary and Tait.(^{27}) 1987</td>
<td>12 house officers; repeated measures across 3 conditions: after night on duty, night off duty, or night admitting emergency cases</td>
<td>Vigilance and cognitive battery (Y)</td>
<td>Large amount of intersubject variability; short-term recall reduced after call night (P&lt;.05)</td>
<td>N</td>
<td>?</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

*Question marks indicate that data were not reported in original article. PASAT indicates Paced Auditory Serial Addition Task; NA, not applicable; PGY, postgraduate year; ABSITE, American Board of Surgery In-Training Examination; ECG, electrocardiogram; IV, intravenous line; and ED, emergency department.*

©2002 American Medical Association. All rights reserved.
attending surgeon on most of the cases. Goldman was likely unblinded to the residents’ sleep condition. In this study, Goldman et al perceived what they termed operative inefficiency, with 30% more surgical time required because of poorly planned maneuvers in 4 of the 5 residents with little sleep (<2 hours of sleep in the last 24 hours). More recently, a large study compared the frequency of significant surgical complications for 3 call schedules relative to the time of surgery.31 Although the investigators concluded that there were no overall differences, complication rates for these surgical residents were 45% higher (P < .02) if the residents had been on call the previous night. Two simulated laparoscopy studies found significantly more errors and observed that more time was required to perform procedures on postcall mornings.35,36

These data suggest that surgical residents may be more vulnerable to performance decrements for skills involving fine motor skills than to tasks requiring cognitive skills. We believe that there are 3 possible explanations for the minimal effect of short-term sleep loss on cognitive performances of surgical residents. First, it is possible that surgical residents self-select for a job with demanding hours and that they are in fact less vulnerable to the effects of sleep loss. The second possibility is that they have achieved a ceiling effect from chronic partial sleep loss, and a further decrement from short-term sleep loss cannot be appreciated. Third, the residents who were most vulnerable to the effects of sleep loss may have selectively refused to participate in these studies.

Several issues will be important to address in future studies. Performances in surgical residents after both short-term and chronic partial sleep loss should be compared with well-rested control performances. The neurobehavioral assays used to evaluate the effects of sleep loss should be validated in other populations to be sensitive to the effects of sleep loss. Ideally, these assays should provide information on the residents’ clinical performances. Future studies should have adequate statistical power to allow conclusions to be drawn from negative findings, should include all eligible residents, and should evaluate performances at several circadian points (or at least at the circadian nadir, 2-9 AM). Finally, stimulant use, including caffeine, should be avoided during study or added to the analysis as a confounding variable.

In summary, significant decrements exist in procedural skills in postcall surgical residents, demonstrating that this group is not immune to the effects of sleep loss. In addition, surgical complication rates may be increased when surgical residents perform surgery after they have been on call. This is an area, therefore, in which further study is indicated.

Nonsurgical Residents

Our MEDLINE search identified 23 peer-reviewed studies evaluating the impact of sleep loss on performances in nonsurgical physicians in training. The majority of studies have characterized the impact of short-term sleep loss on psychomotor function. Several studies have included a simulated medical task as part of the experimental design, sometimes in combination with a number of additional psychomotor tasks. Friedman et al38 studied 14 medical interns’ performances on a 20-minute electrocardiogram interpretation task in both the rested state (mean, 7.0 hours of sleep on the night prior to testing; range, 5.5-8.5 hours) and the sleep-deprived state (mean, 1.8 hours; range, 0-3.8 hours). Efficiency and accuracy of performance were impaired in the sleep-deprived condition; however, factors such as circadian timing and practice effects were not controlled for. In another study, the overall performance of 15 residents in laboratory report interpretations deteriorated after at least 8 hours of cumulative sleep loss (over 1-3 days).39

Not all tests in nonsurgical house staff have shown such performance decrements. In 1 study of 43 pediatric residents that included measurement of performance on both board-type questions and several simulated procedural tasks (performed on cats), including intubation, vein catheterization, and arterial catheterization, after 24 vs 36 hours of continued wakefulness, no significant differences were found on any of the tasks, except for efficiency of arterial cannulation.47 However, circadian timing of the testing, practice effects, differences in accumulated sleep debt, and previous training experiences were potential confounders in this study.

In a more recent study that observed the effect of training experience (first- and second-year residents vs third- and fourth-year), Lingenfelser et al41 examined the performance on a number of psychomotor tasks of 40 residents in the “off-duty” state (≥6 hours of reported sleep the previous night) and after 24 hours on call. Performances on a simulated electrocardiogram, short-term recall of a list of things to do, and reaction times all deteriorated after being on call; these postcall performance deficits were similar for junior and senior residents, suggesting a lack of adaptation over time to the sleep-deprived state. One study of emergency department physicians found that both completion time in a simulated intubation task and clinical accuracy in a triage task were worse for night-shift physicians than for day-shift physicians and that performance deteriorated across night shifts but not across day shifts.48 Performance deterioration across the night shift likely results from both insufficient sleep and circadian rhythm differences. Total sleep times, as measured with polysomnography, for the emergency physicians on night-shift duty were significantly less than sleep times for physicians on day-shift duty.49 The effects of sleep loss on performance of clinical tasks appeared similar to results with simulated tasks. A study of second-year residents in the emergency department found significant postcall reductions in the comprehensiveness of history and physical examination documentation.43 However, a second study in which attending physicians directly observed residents on
a clinical patient evaluation task found no differences in performances of post-call residents compared with residents who had spent the previous night off duty.\textsuperscript{45} Specifically, attending physicians scored the patient-physician interaction (clinical interview and write-up) of medical interns in a standardized patient encounter; no significant differences in performance were found. The small size (n=7) and differences in performance scores for each condition raise issues of the adequacy of statistical power to substantiate negative findings. In another study of clinical performance, Christensen et al\textsuperscript{40} found that radiology residents’ accuracy in detecting pulmonary nodules on a series of radiographs did not vary with rested and nonrested conditions. Accuracy was equally poor (61% correct) in both groups. The prior sleep conditions, however, were not well defined; circadian timing was not controlled for and no information was included about efficiency of performance.

SLEEP LOSS IN RESIDENTS AND PATIENT SAFETY

Large-scale research addressing the prevalence and causes of medical error is relatively new, and most documentation of medical error is neither systematic nor complete.\textsuperscript{80-83} For these reasons, the relative role of insufficient sleep among house staff in preventable adverse events is simply not known. Aside from the surgical complications study cited herein, no studies have examined the effects of resident fatigue on patient safety. The Agency for Healthcare Research and Quality is presently funding more than $300 million of research aimed at identifying causes and solutions for medical error and has included in this budget studies evaluating the direct effect of resident sleep and scheduling on medical error (http://www.npsf.org/html/res_catalog/catalog.html).

EFFECTS OF SLEEP LOSS ON RESIDENT SAFETY

Our search identified 10 peer-reviewed studies evaluating the effects of sleep loss and work schedules on resident health and safety. Training institutions are responsible for the health and well-being of their trainees, as well as that of their patients. A number of studies have documented increased stress and depression,\textsuperscript{84-86} somatic complaints,\textsuperscript{87} and pregnancy-related complications in residents related to sleep loss and fatigue.\textsuperscript{88} However, the greatest documented danger of sleep loss for medical residents is the risk of motor vehicle crashes.\textsuperscript{80-82,93} Several surveys have shown higher motor vehicle crash rates during residency training, particularly during postcall periods.\textsuperscript{80,82,93} While these studies lack definitive controls, the threat to personal safety of residents and interns appears significant. Further research is needed to assess the specific risks associated with sleep loss in physicians in training.

SLEEP INERTIA

Sleep inertia is defined as a clouded sensorium or incomplete arousal from sleep.\textsuperscript{94} Sleep inertia is most likely to occur upon an elicited arousal from a sleep period lasting 3 to 6 hours.\textsuperscript{94,95} Behaviorally, sleep inertia manifests as slowed speech, substantial performance deficits, poor memory, and impaired decision making.\textsuperscript{94} With either time or sufficient stimulation (eg, physical activity or caffeine), sleep inertia is reversed.\textsuperscript{94} Although sleep inertia may result in profound impairments,\textsuperscript{94,95} very little is known about the effects of sleep inertia in residents answering pages or responding to emergencies in the middle of the night.

COUNTERMEASURES FOR SLEEPINESS AND FATIGUE

The American Medical Student Association, the Committee of Interns and Residents, and Public Citizen have recently petitioned the Occupational Health and Safety Administration to reduce work hours to 80 hours per week and on-call time to periods of less than 24 hours continuously (http://www.citizen.org/hrg/publications/1570.htm). Similarly, the Accreditation Council on Graduate Medical Education has recently mandated an 80-hour workweek for all US residents, to take effect in July 2003. This was first proposed by the Bell Commission Report in 1985 and remains controversial.\textsuperscript{96,97} In the absence of evidence for specific work-hour restrictions, however, an optimal schedule for medical trainees has yet to be determined. Moreover, an 80-hour workweek would far exceed the maximum work hours for pilots\textsuperscript{98} and commercial drivers.\textsuperscript{99} Most importantly, reducing work has never been shown to increase sleep in residents. A recent study of night float–protected interns showed that trainees with protected time failed to use the time to sleep and slept the same amount as when on call.\textsuperscript{57} A survey of surgical house staff reported that surgical residents average 5 to 6 hours of sleep across the week whether on every-other-night, every-third-night, or every-fourth-night schedule.\textsuperscript{74} These data suggest that work-hour restrictions alone are not sufficient to counter the effects of sleep deprivations on physician performance.

Central nervous system stimulants have been tested for effectiveness in improving performance following sleep loss. High-dose caffeine, modafinil, and D-amphetamine are effective in reducing sleepiness, as measured with polysomnography, and enhancing vigilance performance, again in individuals following short-term (reduced sleep for <2 days) sleep loss.\textsuperscript{99-103} However, potential health risks with regular use of any of these drugs should preclude consideration as countermeasures for chronic sleep loss in physicians in training, and caffeine does not correct mathematical calculation decrements with sleep loss.\textsuperscript{100} The most effective countermeasure for sleepiness is sleep. A 2- to 8-hour nap prior to 24 hours of sleep loss can improve vigilance and minimize sleepiness for 24 hours.\textsuperscript{104} Naps as short as 15 minutes can significantly ameliorate the performance decrements if provided at 2- to 3-hour intervals during 24 hours of sleep deprivation.\textsuperscript{105} Two-hour naps every 12 hours ameliorate...
SLEEP LOSS AND FATIGUE IN RESIDENCY TRAINING

performance decrements across 88 hours of sleep deprivation.103 Naps must be taken for more than 2 hours to minimize sleep inertia.3 The time of the day most refractory to countermeasures is the circadian nadir, 2 AM to 9 AM.106 Bright light therapy during these hours to activate the circadian system has been studied and may improve vigilance performance.107

CONCLUSION

Achieving an optimal balance between the comprehensiveness of training and adequate sleep or rest time to allow for effective learning and to minimize human error in health care delivery should be an important focus for residency programs. Sleep is the most effective countermeasure for sleep loss, and every intern and resident should understand the importance of coming to work well rested. Napping—as little as 30 minutes every 3 hours—and occasional low-dose caffeine may provide safe countermeasures for prolonged shifts, although their effectiveness remains to be evaluated in rigorous scientific protocols of residents. Efforts should be made to minimize chronic partial sleep loss. Recognizing the inevitable need for physicians to work long hours in some circumstances, steps should be taken to devise optimal work schedules, increase efficiency in performing work duties, develop workstation environments that minimize distractions from patient care, implement systems for detecting high-risk adverse events in patients, and schedule naps during long work shifts. More controlled trials are needed to evaluate the broader effects of sleep loss and fatigue on physicians in training.

Funding/Support: This work was funded in part by National Heart, Lung, and Blood Institute Sleep Academic Award HL 03651.

REFERENCES

4. Dement WC, Miller NM. It’s time to wake up to the importance of sleep disorders. JAMA. 1993;269:1548-1550.
17. Kline S, Gally, Paradoxical sleep deprivation applied two days after the end of training retards learning. Physiol Behav. 1988;43:213-216.
25. Franco P, Szwilkowski H, Darmais M, Kahn A. Influence of ambient temperature on sleep character-
SLEEP LOSS AND FATIGUE IN RESIDENCY TRAINING


82. Geer RT, Jobes DR, Tew JD, Stephens LH. Incidence of automobile accidents involving anesthesia residents after on-call duty cycles. Anesthesiology. 1997;87:493A.


