# The Effects of Sleep Loss on Cognitive Performance of Resident Physicians

C. H. M. Jacques, MD, PhD, James C. Lynch, MS, and Judith S. Samkoff, MD, ScM
*Hershey, Pennsylvania*

*Although the long hours worked by resident physicians have raised concern in the public, legislators, and members of the medical profession, the consequences of sleep loss on the ability of residents to perform in clinical settings is unknown. The present study examined the effects of various amounts of reported sleep on cognitive performance measured by the American Board of Family Practice in-training examination. A total of 353 family practice residents in 21 programs who took the examination in Pennsylvania in 1988 were studied. Linear regression analysis demonstrated a statistically significant ($P < .05$) decline in composite test score with decreasing sleep on the night before the examination for residents in each year of training. Loss of one night's sleep resulted in changes in test scores that were approximately equivalent in magnitude to the change that occurred in test scores between residents in the first and third year of training. The results suggest that prolonged testing over several hours may be necessary to detect the subtle but significant differences in cognitive performance that are present with relatively mild degrees of sleep loss.* **J FAM PRACT 1990; 30:223-229.**

The death of a young woman in a New York teaching hospital in 1984 and the subsequent grand jury investigation and national media coverage suggesting that resident fatigue and lack of supervision may have contributed to inadequate care have set in motion forces that will substantially alter residency training.[1,2] The State of New York recently implemented legislation to limit resident working hours, and several other states have proposed similar legislation.[3] The American Association of Medical Colleges has recommended guidelines for resident hours.[4] Major changes limiting resident hours have already occurred for some disciplines under the direction of the Accreditation Council for Graduate Medical Education.[5]

Recent reviews, editorials, and commentaries have cited the long working hours and sleep deprivation enforced by residency call schedules as major factors contributing to resident stress and fatigue.[6–10] Residents have expressed concern about the quality of care they provide when they are fatigued.[11–13] The AMA Center for Health Policy Research reported that the average resident spent 74.2 h/wk performing residency-related duties in 1987, exclusive of moonlighting. Interns averaged 85 h/wk for all specialties and over 100 h/wk for surgical specialties. Comparison of the data between similar surveys in 1983 and 1987 showed an increase in the hours worked by residents, caused primarily by an increase in the number of hours of patient care.[14]

Despite legislative action and concern by many parties, relatively little is known about the effects of either acute or chronic sleep deprivation on the performance of residents. In one of the most widely quoted articles on sleep loss among residents, Friedman et al[15] in 1971 noted that residents who had been up the night before made significantly more errors than rested residents on reading a standardized electrocardiogram. Since that time 20 reports regarding some aspects of fatigue or sleep deprivation in residents have appeared in the literature.[16–35] The studies consistently show that mood, attitude, and perceived efficiency deteriorate with sleep loss.[16,21,22,29,30,33]

Reports on measures of performance, however, have not been consistent. Although many studies have shown a deterioration in some measure of performance after the loss of a single night's sleep,[15,17,18,21,23,24,26,27,30,33,34] other

*Submitted September 29, 1989.*

*From the Department of Family and Community Medicine and the Department of Biostatistics, College of Medicine, The Milton S. Hershey Medical Center, The Pennsylvania State University, Hershey, Pennsylvania. Requests for reprints should be addressed to C. H. M. Jacques, MD, PhD, Department of Family and Community Medicine, PO Box 850, The Milton S. Hershey Medical Center, Hershey, PA 17033.*

studies, in some cases using the same tests, have shown no change in performance with sleep loss.[19,22,28,29,31,32,35] The inconsistencies of the results reflect the generally small numbers of residents studied, the inconsistent definition of sleep loss, and the wide variety of tests used to measure performance. Many of the studies on performance, especially those in the past 10 years, have used short psychometric tests. These studies have been criticized both on methodologic grounds and for the insensitivity of psychometric tests in measuring sustained cognitive performance in clinical situations.[36,37]

The purpose of the present study was to use a validated, standardized test of cognitive performance on a large sample of resident physicians exposed to various degrees of sleep loss the night before the examination. A test sufficiently long to measure subtle changes in sustained performance was desirable. Each year the American Board of Family Practice provides accredited residency programs in family practice with a standardized 4-hour in-training examination developed to assess the cognitive knowledge of the residents. The present study examined the effects of reported sleep loss on in-training examination scores for 353 residents from 21 family practice residencies who took the examination in the state of Pennsylvania in 1988.

## METHODS

### In-Training Examination

The American Board of Family Practice In-Training Examination is a written test given on the second Friday of November to residents of approved family practice residencies. The examination is conscientiously proctored and administered in a standardized fashion to ensure that national norms will be valid and that the programs will know that all results were obtained under equal conditions. The examination includes questions from the following subject areas: internal medicine, surgery, obstetrics, community medicine, pediatrics, psychiatry and behavioral sciences, gerontology, and gynecology. These areas are the same as those covered on the certification and recertification examinations for the American Board of Family Practice.

The examination is divided into two parts. Book I consists of a 3-hour test of various types of multiple-choice questions designed to assess clinical knowledge. Matching questions, one-best-response questions, and true-false questions are used. Some questions require the resident to look at a photograph or illustration in order to answer the question. Book II is a 1-hour section that contains only clinical set problems. These questions provide clinical material on which the examinee must base management decisions. The questions follow the format of providing some introductory information about a patient and then asking the resident to answer questions relating to the diagnosis, treatment, and management of the patient. Additional information concerning laboratory results, treatment outcomes, and so on are then given, and additional questions based on the new information are then asked.

The residents' raw scores are standardized across the total group of residents for one set of standardized scores for each of the eight major disciplines. Standardized scores range from 200 to 800 with a mean of 500 and a standard deviation of 100. A similar standardized score is provided for the clinical set problems. The composite test score is derived from the weighted average of the standard scores for the Book I multiple-choice questions, the Book I true-false questions and the Book II clinical set problems. All scores in each category are also reported as total percentile ranking within the group of all residents taking the examination and percentile ranking within each year of residency.

### Self-Reported Sleep Data

Information regarding the total number of hours slept the night before the examination, total number of hours worked per week, and the average on-call frequency was obtained by asking residents to fill out a questionnaire at the time of administration of the in-training examination. The examination proctors were instructed to distribute the questionnaire in a way that did not interfere with the testing protocols required by the American Board of Family Practice. Examination results were provided to program directors in January 1989. Each program director participating in the study was asked to release the composite score and the total clinical set problems score for each resident taking the examination using a code number that did not identify individual residents or programs.

The in-training examination was conducted at 30 Pennsylvania family practice residency programs ranging in size from 12 to 33 residents in November of 1988. All programs were invited to participate in this study. Twenty-one programs provided complete questionnaire data and examination score data representing a total of 353 residents.

### Statistical Methods

The relationships of sleep, on-call frequency, and hours worked per week to test scores were examined using linear regression. Because the between-site variance and sleep experience differed across year of training, each year was modeled separately. Regression models were fit

**TABLE 1. SELF-REPORTED DATA FROM FAMILY PRACTICE RESIDENTS (N = 353) IN PENNSYLVANIA**

| Year | No. | Total Sleep (hours) | Work Week (hours) | On-Call Frequency (days) |
|---|---|---|---|---|
| 1 | 115 | 5.7 ± 2.0 | 88.1 ± 10.0 | 4.0 ± 0.6 |
| 2 | 127 | 6.3 ± 1.7 | 77.2 ± 12.3 | 5.5 ± 1.6 |
| 3 | 111 | 6.3 ± 1.3 | 67.9 ± 12.6 | 7.1 ± 3.0 |
| All data | 353 | 6.1 ± 1.7 | 77.8 ± 14.2 | 5.5 ± 2.3 |

by maximum likelihood estimation using program 3V in BMDP statistical software.[38] The significance of the linear relationship between sleep and test score was assessed using the *t* statistic associated with the slope coefficient. Estimated slopes of the linear regression are reported as ±1 SD.

The data were collected as a cluster sample from 21 sites instead of by simple random sampling of family practice residents in Pennsylvania. To account for cluster sampling rather than random sampling, random site effects were included in the regression model.[39] The sleep-adjusted mean test scores were compared using a Bonferroni correction for multiple comparisons. Likelihood ratio tests were used to determine the effect of on-call frequency and total hours worked per week on test performance. All tabulated values are reported as mean values ±1 SD.

## RESULTS

The self-reported data from all 353 residents for the average hours sleep on the night before the examination, average hours worked per week, and on-call frequency are presented in Table 1. Information on the number of hours worked per week and the total number of hours of sleep on the night before the in-training examination is shown graphically in Figure 1 and Figure 2, respectively.

Family practice residents in Pennsylvania reported working an average of 77.8 h/wk for all residents, with first-year residents working 88.1 h/wk. These figures are similar to the 70.5 hours for all family practice residents and 82.3 hours for first-year residents reported in a national survey.[14] Over one half of the first-year residents and about one third of the second-year residents in this study reported working more than the 80 h/wk recommended by the American Association of Medical Colleges. Mean on-call frequency for first-year family practice residents in Pennsylvania is every fourth night and decreases to about once per week by the third year, with considerable variation between programs. The mean number of hours of sleep for the study population on the night before the in-training examination was 6.1 hours. The percentage of residents with fewer than 4 hours sleep for the first, second, and third year of training were 13.9%, 10.2%, and 4.5%, respectively.

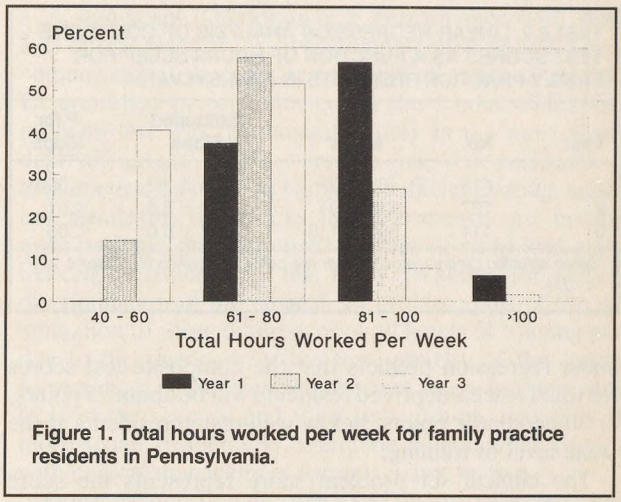


**Figure 1. Total hours worked per week for family practice residents in Pennsylvania.**

The results of simple linear regression analysis of the effect of sleep the night before the in-training examination on the composite test score are shown in Table 2 and graphically represented in Figure 3. As expected, test scores are statistically different for each year of training, presumably reflecting increased accumulation of knowledge with training. Also as expected, a large variation in test scores occurs at each sleep state, representing the wide range of abilities of the test takers. Nonetheless, a statistically significant decline in composite test score occurs with decreasing sleep for each year of training. The strongest statistical significance is in the first year. The

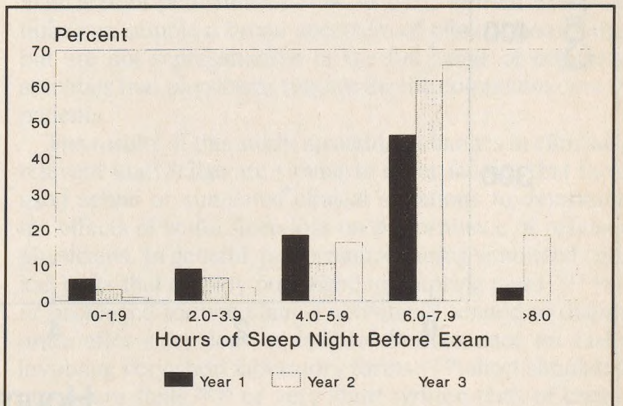


**Figure 2. Total hours of sleep on the night before the in-training examination for family practice residents in Pennsylvania.**

**TABLE 2. LINEAR REGRESSION ANALYSIS OF COMPOSITE TEST SCORES AS A FUNCTION OF HOURS SLEEP FOR FAMILY PRACTICE RESIDENTS IN PENNSYLVANIA**

| Year | No. | Mean* | Estimated Slope | *P* for Slope |
|---|---|---|---|---|
| 1 | 115 | 485 ± 71 | 8.4 ± 3.2 | <.0001 |
| 2 | 127 | 514 ± 94 | 9.4 ± 3.5 | <.01 |
| 3 | 111 | 550 ± 116 | 11.7 ± 5.0 | <.05 |

**Sleep adjusted mean scores by year are pairwise significantly different, all P < .001.*

**TABLE 3. LINEAR REGRESSION ANALYSIS OF CLINICAL SET PROBLEM TEST SCORES AS A FUNCTION OF HOURS SLEEP FOR FAMILY PRACTICE RESIDENTS IN PENNSYLVANIA**

| Year | No. | Mean* | Estimated Slope | *P* for Slope |
|---|---|---|---|---|
| 1 | 115 | 478 ± 61 | 13.8 ± 3.9 | <.001 |
| 2 | 127 | 511 ± 186 | 6.7 ± 4.3 | .13 |
| 3 | 111 | 552 ± 147 | 6.2 ± 5.8 | .29 |

**Sleep adjusted mean scores by year are pairwise significantly different, all P < .001.*

linear regression predicts that the composite test scores for totally sleep-deprived residents will be about 75 points, or 30 percentile points, below well-rested residents at the same level of training.

The clinical set problem score represents the score from the final 1 hour of testing and is a component of the entire 4-hour composite test score. The results of simple linear regression of clinical set problem scores as a function of sleep are shown in Table 3. Although there is a trend for clinical set problem scores to decline with decreasing sleep in the second and third years, a statistically

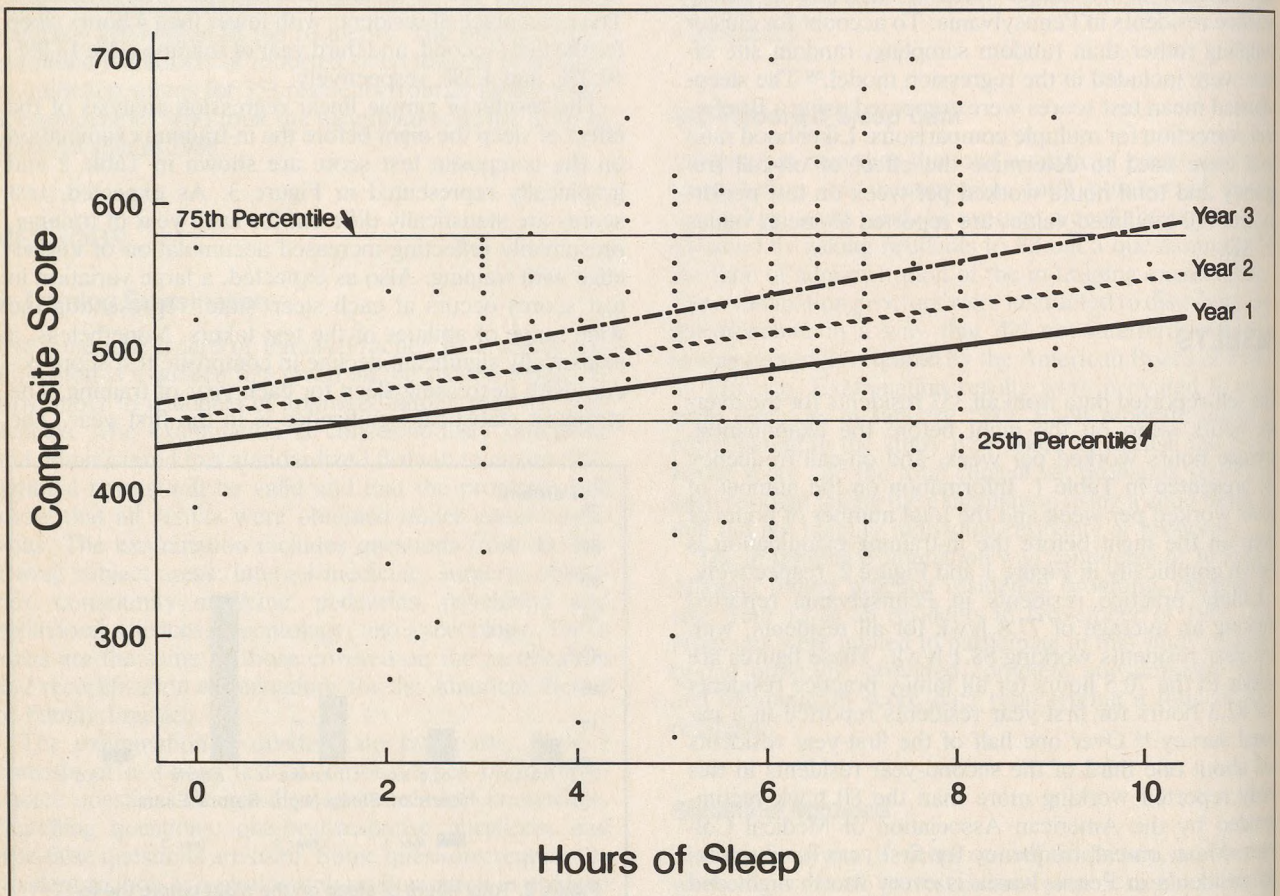


**Figure 3. Linear regression of composite test score as a function of hours sleep on the night before the in-training examination for family practice residents in Pennsylvania.**

significant relationship is only noted in the first year of training, where the greatest number of sleep-deprived subjects were present.

Multiple linear regression analysis was performed for on-call frequency and total hours worked per week to determine whether these variables were additional predictors of test performance. On-call frequency and total hours worked did not contribute significantly to the overall results of the composite test scores or the clinical set problem test scores after adjusting for hours of sleep.

## DISCUSSION

Most of the 353 residents involved in this study were not severely sleep deprived and were not involved in the exhausting work schedules characteristic of some residency programs. Nonetheless, a significant correlation was noted between test scores for the in-training examination and the amount of sleep the night before the test. The regression analysis of the composite test score as a function of sleep demonstrated a 30 percentile point change in the score between well-rested and sleep-deprived residents at the same level of training. Put another way, loss of one night's sleep results in a change in test score approximately equivalent to the change that occurs between the first and third year of residency training.

The degree to which cognitive tasks are sensitive to sleep loss is thought to be determined by a number of task-specific factors as well as environmental and psychological factors.[40,41] Early studies by Wilkinson found that after moderate sleep deprivation of one night, significant performance impairment on serial reaction time and vigilance tasks could be measured only during the last half of these 30-minute tasks.[42,43] Such results led Wilkinson to conclude that for a test to be sensitive to moderate sleep deprivation, it must be prolonged to at least 30 minutes and preferably be 1 hour in duration. Donnell,[44] using a modified addition task, found that after 32 hours of sleep deprivation, accuracy of performance did not differ from baseline until 50 minutes of testing had elapsed. More recent studies have demonstrated that shorter duration tasks such as a 10-minute auditory reaction time task[45] or a 10-minute reaction time task[46] are sensitive to continuous sleep loss of one night's duration. Even shorter tests can demonstrate differences in performance with extended periods of sleep loss, especially if subjects are repeatedly tested under intense workload conditions.[41] Performance of sleep-deprived subjects can be substantially improved by increasing mental effort, particularly if the task is interesting or if the stimulus is strong enough. Tasks that are prolonged, dull, or repetitive have been shown to be most sensitive to sleep deprivation.[47]

A significant deterioration in performance with sleep loss of one night or less was demonstrated on the 4-hour composite test score for residents in all years of training. Deterioration in performance on the 1-hour clinical set problem test was demonstrated only in the most sleep-deprived group, which were the first-year residents. It appears that if cognitive testing is sufficiently long, subtle but significant differences in performance are present even with the relatively mild degrees of sleep loss experienced by residents in this study. Whether the lack of discrimination of the clinical set problem part of the examination to show differences at all levels of training was due to the shortness of the test, novelty of the testing format, lack of sufficient numbers of sleep-deprived subjects in the second- and third-year groups, or ability of more senior residents to overcome fatigue when dealing with clinical management formats is not known.

One must question the validity of tests involving written questions in predicting performance in actual clinical situations. In an excellent review on the subject, Neufeld and Norman[48] have concluded that although written tests are the most common form of assessing clinical competence, available evidence is inconclusive. The cognitive aspects of clinical competence appear to have two main components, a knowledge factor and a performance factor. Well-constructed written questions based on clinically relevant material measure the knowledge factor, as demonstrated by the increase in test scores with increasing year of training in the current study. There is some evidence that performance on clinically relevant multiple-choice question examinations is moderately and positively correlated with well-constructed simulated clinical encounters.[48] This finding, however, is not consistent. Several studies have shown poor correlation between written test scores on in-training examinations and other measures of performance.[49–51] At best, written examinations can sample a broad spectrum of clinical knowledge but are not representative of the full range of cognitive activities that physicians require for the competent care of patients.

The results of this study measuring changes in clinically relevant knowledge are similar to other studies that have used actual or simulated clinical situations to determine the effects of acute sleep loss on performance of resident physicians. In general, performance during simulated clinical tasks that employ prolonged monitoring tasks[15,17,18,21] or prolonged routine clinical activities[34] tended to deteriorate after sleep loss or fatigue. Performance on tasks involving very short laboratory forms,[20,28] short simulated procedure skills,[28,35] or very short written tests of cognitive knowledge[28,35] tended to show little or no deterioration with sleep loss, although performance of some procedural skills required more time. Two studies with untimed tasks, one involving actors as patients[29] and one

involving a radiological task,[19] also failed to show changes with fatigue or sleep loss. Although simple conclusions are difficult to draw because of the diversity of experimental conditions, it appears that prolonged testing of routine tasks is necessary to detect changes in performance in mildly to moderately sleep-deprived residents.

Several recent studies with surgical residents exposed to long working hours and frequent night call have failed to demonstrate statistically significant differences between the performance of experimental groups with acute sleep deprivation and control groups.[28,31,32] Total testing time in each of these studies was 1 hour or less, and some of the psychometric tests used for performance evaluation were as short as 30 seconds. These studies have been criticized because the long work hours and every-other-night call schedule of the study groups may have resulted in evaluating two more or less equally exhausted populations. In the present study using more rested family practice residents, multiple regression analysis failed to show a significant effect of call frequency or total hours worked per week on in-training examination test scores. Better studies need to be designed specifically to separate the effects of acute and chronic sleep deprivation on performance, especially when determining the effects of sleep loss on performance in residency programs in which long hours and frequent call are common.

In conclusion, numerous studies have demonstrated that scales of anger, hostility, and depression increase for interns and residents after sleep loss imposed by night call. The results of this study demonstrate that cognitive performance measured by a well-constructed written examination of clinical information significantly deteriorates after sleep loss the night before the examination. Although the effects of acute and chronic sleep deprivation on actual clinical care have yet to be directly measured in large well-controlled studies, the cumulative evidence collected to date strongly suggests that both the mood and performance of resident physicians are adversely affected by the moderate sleep loss imposed by typical resident work schedules.

## Acknowledgment

Support of this research was provided by grant funding from the US Department of Health and Human Services, grant No. PE13013. This research could not have been accomplished without the support of the program directors and staff of the 21 family practice residency programs in Pennsylvania who administered the questionnaire and provided the examination scores.

## References

1. McCall T: No turning back: A blueprint for residency reform. JAMA 1989; 261:909–910
2. Asch DA, Parker RM: The Libby Zion case: One step forward or two steps backward. N Engl J Med 1988; 318:771–778
3. Sandrick K: Proposals regulating resident work hours stall. Hospitals 1988; 62 (April):66–68
4. Executive Council AAMC: Resident supervision and hours: Recommendations of the Association of American Medical Colleges. J Med Educ 1988; 63:417–426
5. Page L: Too many hours, too little sleep? Focus turns to residents' schedules. Am Med News, January 1989, pp 6–7
6. Butterfield P: The stress of residency—A review of the literature. Arch Intern Med 1988; 148:1428–1435
7. Squires B: Fatigue and stress in medical students, interns, and residents: It's time to act. Can Med Assoc J 1989; 140:18–19
8. Levin R: Beyond the men of steel—The origins and significance of house staff training stress. Gen Hosp Psychol 1988; 10:114–121
9. Dilts PU, Dilts SL: Stress in residency: Proposals for solution. OB/GYN 1987; 157:1093–1096
10. Colford JM, McPhee SJ: The ravelled sleeve of care—Managing the stresses of residency training. JAMA 1989; 261:889–893
11. Wilkinson RT, Unrey CA: Duty hours of young hospital doctors: Effects on the quality of work. J Occup Psychol 1975; 48:219–229
12. McManus IC, Lockwood DNJ, Cruickshank JK: The preregistration year: Chaos by consensus. Lancet 1977; 1:413–416
13. Lewittes LR, Marshall VW: Fatigue and concerns about quality of care among Ontario interns and residents. Can Med Assoc J 1989; 140:21–24
14. Silberger AB, Than SL, Marder WD: The changing environment of resident physicians. Health Aff 1988 (suppl):121–133
15. Friedman, RC, Bigger JT, Kornfeld DS: The intern and sleep loss. N Engl J Med 1977; 285:201–203
16. Friedman, RD, Kornfeld, DS, Bigger TJ: Psychological problems associated with sleep deprivation in interns. J Med Educ 1973; 48:436–441
17. Goldman LI, McDonough MT, Rosemond GP: Stresses affecting surgical performance and learning: I. Correlation of heart rate, electrocardiogram, and operation simultaneously recorded on videotapes. J Surg Res 1972; 12:83–86
18. Beatty J, Ahern BK, Katz R: Sleep Deprivation and the Vigilance of Anesthesiologists During Simulated Surgery. Vigilance Theory, Operational Performance, and Physiological Correlates. New York, Plum Press, 1977, pp 511–527.
19. Christenson ET, Dietz GW, Murry RC, Moore JB: The effect of fatigue on resident performance. Radiology 1977; 125:103–105
20. Poulton EC, Hunt GM, Carpenter A, Edwards RS: The performance of junior hospital doctors following reduced sleep and long hours of work. Ergonomics 1978; 21:279–295
21. Leighton K, Livingston M: Fatigue in doctors, letter. Lancet 1983; 1:1280.
22. Ford CV, Wentz DK: The internship year: A study of sleep, mood states, and psychophysiological parameters. South Med J 1984; 77:1435–1442
23. Harrah CH Jr: Effects of Sleep Deprivation on Neuroposychological Functioning of Medical and Surgical Residents, Dissertation, Pasadena, Calif, Fuller Theological Seminary, 1984
24. Hawkins MR, Vichick DA, Silsby HD, et al: Sleep and nutritional deprivation and performance of house officers. J Med Educ 1985; 60:530–535
25. Klose KJ, Wallace-Barnhill GL, Craythorne NWB: Performance test results for anesthesia residents over a five-day week including on-call duty. Anesthesiology 1985; 63(3A):A485
26. Narang V, Laycock JRD: Psychomotor testing of on-call anesthesiologists. Anesthesiology 1986; 41:868–869
27. Denisco RA, Drummond JN, Gravenstein JS: The effect of fatigue on the performance of a simulated anesthetic monitoring task. J Clin Monit 1987; 3:22–24

28. Reznick RK, Folse JR: Effect of sleep deprivation on the performance of surgical residents. Am J Surg 1987; 154:520–525
29. Engel W, Seime R, Powell V, D'Allesandri RD: Clinical performance of interns after being on call. South Med J 1987; 80:761–763
30. Hart RP, Buchsbaum DG, Wade JB, et al: Effect of sleep deprivation on first-year residents' response times, memory and mood. J Med Educ 1987; 62:940–942
31. Bartle EJ, Sun JH, Thompson L, et al: The effects of acute sleep deprivation during residency training. Surgery 1988; 104:311–316
32. Deaconson TF, O'Hair DP, Levy INF, et al: Sleep deprivation and resident performance. JAMA 1988; 260:1721–1727
33. Sharp KH, Vaughn GM, Cosby PW, et al: Alterations of temperature, sleepiness, mood, and performance in residents are not associated with changes in sulfoxymelatonin excretion. J Pineal Res 1988; 5:499–512
34. Bertram DA: Characteristics of shifts and second-year resident performance in an emergency department. NY State J Med 1988; 88:10–14
35. Storer JS, Floyd HH, Gill WL, et al: Effects of sleep deprivation on cognitive ability and skills of pediatric residents. Acad Med 1989; 64:29–32
36. Parker JBR: The effects of fatigue on physician performance—An underestimated cause of physician impairment and increased patient risk. Can J Anaesth 1987; 34:489–495
37. McCall T, Swartz CE, Ritchee K, et al: Sleep deprivation and performance of residents, letter. JAMA 1989; 261:859–863
38. Dixon WJ (chief ed): BMDP Statistical Software. Berkeley, University of California Press, 1985.
39. Cochran WG: Sampling Techniques. ed 3. New York, John Wiley & Sons, 1977
40. Johnston LC: Sleep Deprivation and Performance. Biological Rhythms, Sleep and Performance. New York, John Wiley & Sons 1982; pp 111–141
41. Heslegrave RJ, Angus RG: The effects of task duration and work-session location on performance degradation induced by sleep loss and sustained cognitive work. Behav Res Meth Ins Comp 1985; 17:592–603
42. Wilkinson RT: Interaction of lack of sleep with knowledge of results, repeated testing, and individual differences. J Exp Psychol 1961; 62:263–271
43. Wilkinson RT: Effects of up to 60 hours of sleep deprivation on different types of work. Ergonomics 1964; 7:175–186
44. Donnell JM: Performance decrement as a function of total sleep loss and task duration. Percep Motor Skills 1969; 29:711–714
45. Lisper HD, Kjellberg A: Effects of 29 hour sleep deprivation on rate of decrement of a 10 minute activity reaction time task. J Exp Psychol 1972; 96:287–290
46. Glenville M, Broughton R, Wine AM, Wilkinson RT: Effects of sleep deprivation on short duration performance tests compared to the Wilkinson auditory vigrance task. Sleep 1978; 1:169–176
47. Horne JA: Sleep function, with particular reference to sleep deprivation. Ann Clin Res 1985; 17:199–208
48. Neufeld UR, Norman GR: Assessing Clinical Competence. New York, Springer 1985, pp 94–118
49. Borlase BC, Bartle EJ, Moore EE: Does the In-Training Examination Correlate With Clinical Performance in Surgery. How to Evaluate Residents. Chicago, American Board of Medical Specialties, 1986, pp 183–187
50. McLeskey CH, Ward RJ: Validity of written examinations. Anesthesiology 1978; 49:224
51. Skakun EN, Wilson DR, Taylor WC: A comparison of performance on multiple choice tests, oral examinations, and in-training evaluation efforts. Ann R Coll Surg Can 1978; 11:284–285