

# Sleep Loss and Performance in Residents and Nonphysicians: A Meta-Analytic Examination

Ingrid Philibert, MHA, MBA

Accreditation Council for Graduate Medical Education, Chicago, IL

**Study Objectives:** To explore the effect of sleep loss on cognitive function, memory, and vigilance in resident physicians and nonphysicians and on residents' clinical performance.

**Design:** Meta-analysis of 60 studies on the effect of sleep deprivation, with a total sample of 959 physicians and 1,028 nonphysicians and 5,295 individual effect indexes.

**Outcome Measures:** Cognitive performance and performance on clinical tasks under acute and partial chronic sleep deprivation. Additional analyses stratified the data by physician/nonphysician, type of performance, and length and type of sleep loss and assessed the combined effect of several of these factors.

**Results:** Sleep loss of less than 30 hours reduced physicians' overall performance by nearly 1 standard deviation and clinical performance by more than 1.5 standard deviations. The effect of sleep deprivation was larger in nonphysicians than in physicians (corrected *d* value -.995 vs -.880), with

these smaller effects likely resulting from "study factors," primarily variation in the hours without sleep prior and chronically reduced sleep in the "rested" controls in physician studies.

**Conclusions:** The weekly hours and continuous wakefulness permitted under the current national minimum standards for residents may not completely guard against the negative effect of sleep loss on cognitive and clinical performance. Research is needed to explore the effect of continuous duty periods and chronic partial sleep loss in residents and to assess the clinical and educational consequences of sleep loss. The goal should be to combine scientifically based duty-hour limits with broader efforts to enhance patient safety and resident learning.

**Keywords:** Sleep deprivation, cognitive performance, resident physicians, work hour limits, meta-analysis

**Citation:** Philibert I. Sleep loss and performance in residents and nonphysicians: a meta-analytic examination. *SLEEP* 2005;28(11): 1392-1402.

## INTRODUCTION

LONG HOURS AND WORKING OVERNIGHT ARE COMMON IN MANY PROFESSIONS. THEY ARE A PROMINENT FEATURE OF MANY MEDICAL RESIDENCIES, the 3 to 7 years of clinical education in a specialty that follows graduation from medical school. Physicians' work depends on cognitive skills for patient care, and being awake and alert is equally important to the learning process. In 1971, Friedman et al found that residents who had been on call the night before made more errors in reading an electrocardiogram than their rested colleagues.<sup>1</sup> In the ensuing 30 years, many editorials and articles have discussed the effect of sleep deprivation in residents on patient safety and residents' learning and well-being. The reduction in performance after 24 hours of sleep loss has been equated to the effect of a .1% blood alcohol concentration.<sup>2,3</sup> This comparison has received much attention because a physician with a blood alcohol level of .1% would be considered unfit for duty.<sup>4</sup> A growing focus on the relationship between sleep deprivation and reduced performance contributed to legislative proposals to limit resident hours.<sup>5</sup> Reasons cited include risk for medical errors and negative outcomes for residents, including depression, motor vehicle accidents and

needle-stick injuries.<sup>6-9</sup> In July 2003, the Accreditation Council for Graduate Medical Education (ACGME) instituted minimum duty-hour standards for approximately 100,000 residents training in accredited programs.<sup>10</sup> The common minimum standards, shown in Table 1, encompass an 80-hour weekly limit and a limit of 24 hours on continuous duty with an added period of up to 6 hours for the transfer of care and didactic activities. When the ACGME developed these limits, it established the weekly limit to reduce chronic partial sleep loss and the limit on continuous duty (24 plus up to 6 hours) to address acute sleep deprivation.<sup>11</sup>

The ACGME began to set standards for resident hours as far back as the 1980s, but initial steps were followed by extended debate within the educational community about the merits of such limits.<sup>12-14</sup> The establishment of common minimum standards in 2003 was prompted in part by the impending threat of federal regulation. Given the medical profession's reliance on scientific evidence, a factor in the slow progress toward national duty hour limits for residents appears to have been disagreement among

**Table 1**—The ACGME Common Standards for Resident Duty Hours

- An 80-hour limit per week, averaged over 4 weeks
- 1 day in 7 free from all program responsibilities
- Adequate rest must be provided between duty periods (this should generally be 10 hours)
- In-house call no more often than every third night
- A limit on continuous duty of 24 hours+ up to an added 6 hours for transfer of care and didactics; no new patients are allowed after 24 hours
- In-hospital hours during call from home counted
- Moonlighting must be approved by program director
- In-house moonlighting counts toward the weekly 80-hour limit

Accreditation Council on Graduate Medical Education (ACGME) Common Duty Hour Standards, Effective July 1, 2003.<sup>10</sup>

## Disclosure Statement

This was not an industry supported study. Ms. Philibert works for the Accreditation Council for Graduate Medical Education, the organization that developed and enforces the duty hour standards for resident physicians.

Submitted for publication July 2004

Accepted for publication July 2005

Address correspondence to: Ingrid Philibert MHA, MBA, Director of Field Activities, Accreditation Council for Graduate Medical Education, 515 North State Street, Suite 2000; Chicago, IL 60610; Tel: (312) 755-5003; Fax: (312) 755-7498; E-mail: iphilibert@acgme.org

studies of how sleep loss affects performance. An early study concluded that sleep deprivation of less than 46 hours had a minimal effect on human performance.<sup>15</sup> Some studies of sleep deprivation in physicians have found that it impairs performance, others have shown no effect, and, in a few, the sleep-deprived group appear to perform better than the rested controls.<sup>16-17</sup> When the literature produces contradictory findings that frustrates the formulation of scientific theory, meta-analysis can be useful. It facilitates the aggregation data across many studies, by transforming data into *d* values that provide a standardized measure of the difference between the performance of the sleep-deprived and the rested cohorts. Meta-analysis also makes it possible to identify and address moderators and study factors that contribute to contradicting findings.<sup>18-21</sup> The first goal of this meta-analysis was to offer a quantitative counterpart to reviews of the effect of sleep loss on physician performance; a second goal was to compare the findings in physicians and nonphysicians to explore differences found in earlier reviews; a third goal was to aggregate data on clinical performance; and a fourth was to analyze variables that moderate the effect, including type of performance and time without sleep, and differences between laboratory studies and the field studies and “natural experiments” comprising most studies of resident physicians.

## Literature Review

The literature on sleep loss and performance in physicians encompasses tests of cognitive performance and performance on clinical tasks, such as the early work by Friedman et al.<sup>1</sup> They have found that sleep deprivation has a negative effect on human performance relevant to physicians’ work, including cognitive function, working memory, vigilance, fine motor skills and mood.<sup>1,16,17,22-27</sup> Studies of the clinical implications of sleep loss have found that it is associated with a greater complication rate and increased errors and lower effectiveness on actual and simulated care tasks.<sup>28-36</sup> A meta-analysis assessed the effects of duration of sleep loss and type of performance, finding the largest reductions when sleep loss exceeded 46 hours and the pace of performance was dictated by the task.<sup>37</sup> Another meta-analysis found that sleep loss reduced cognitive performance by -1.55 standard deviations in a sample that included physicians and nonphysicians and had a negative effect on mood that approached 3 standard deviations.<sup>38</sup> To the author’s knowledge, no other meta-analyses of the effect of sleep loss on performance in normal adults have been published.

Overnight call is common in many residency programs. It results in acute sleep deprivation, not infrequently approaching 24 hours without sleep. Two comprehensive qualitative reviews of the effect of sleep loss in resident physicians found that sleep loss consistently resulted in reduced cognitive performance, concentration, and mood.<sup>39,40</sup> Both reported a decline in performance after 24 hours without sleep. Studies have shown that alertness is higher and errors are lower if continuous wakefulness is limited to 16 hours.<sup>41-43</sup> A laboratory study in nonphysicians concluded that alertness was compromised once cumulative wakefulness exceeds 15.8 hours.<sup>44</sup> A review of sleep and performance in residents also highlighted the effect of chronic partial sleep loss.<sup>40</sup> Chronic partial sleep deprivation is defined as sleep duration of less than 5 to 6 hours for several consecutive nights.<sup>45</sup> Chronic partial sleep loss is common in residency, with 20% of residents reporting sleep of 5 hours or less per night, and 66% indicating that they sleep 6

hours or less.<sup>46</sup> Residents who report sleeping 5 hours or less are more likely to report having worked in an “impaired condition” and having made medical errors.<sup>46</sup> Laboratory-based studies have found that chronic partial sleep loss has a negative effect on cognitive performance.<sup>45,47-51</sup> Research conducted prior to implementation of the ACGME’s duty-hour standards showed that residents felt that the sleep loss they experienced was detrimental to their performance and learning and affected their well-being.<sup>52</sup>

Moderator variables may contribute to conflicting findings across studies. Moderators explored in prior meta-analyses of sleep and performance have included cognitive versus motor performance versus mood; the extent of sleep loss; and task duration and task complexity. Qualitative reviews have added subject (physician vs nonphysician), and study quality (laboratory vs field studies, with lower control on the study variables in field studies) as additional potential moderators.<sup>39,40</sup> Most studies in nonphysicians are controlled laboratory experiments, while the common approach for studying residents involves field studies that compare the performance of postcall residents to a “rested” resident group or study the same residents in a sleep-deprived and a “rested” state, with far less ability to control the independent variables than is present in laboratory studies. Reduced controls in field studies include variability in the number of hours subjects went without sleep before being tested and chronically reduced sleep in the “rested” residents that serve as controls. Additional differences may include a lack of objective or subjective assessments of sleepiness at time of testing<sup>53</sup> and absence of pretest practice to reduce “practice” or “learning” effects that may mask the effect of sleep deprivation, especially if tests use infrequently practiced activities or have readily discernible solutions.<sup>54</sup> Practice effects have been documented for neuropsychometric tests, including those used to assess the effects of sleep loss.<sup>55,56</sup>

Type of performance is important because mood and vigilance appear to be affected most by sleep loss, while other aspects of cognitive performance are less affected and gross motor performance is quite resilient.<sup>37,38,40,57</sup> Hours without sleep influences the effect, with long-term total sleep deprivation having the most pronounced negative effect.<sup>37,38</sup> Length and complexity of the test or task, and whether it is self-paced or performed at forced speed, also influence the effect.<sup>37,38,58</sup> Finally, interindividual variation in how sleep loss affects performance may be a moderator. Some individuals appear to be profoundly and others minimally affected by the same number of hours without sleep.<sup>59-61</sup>

## Methodology

### Selection and Coding of Studies

Identification of studies started with an electronic search of the literature beginning in 1971, the year a study of sleep loss and performance in resident physicians first appeared in a United States medical or scientific journal. The literature search encompassed studies of sleep loss in physicians and nonphysicians. Databases searched included Medline, PsycLIT, SOCIOFILE, EBSCO-HOST Academic Search Elite, and Web of Science. The reference sections of documents were reviewed for works not found in the electronic records. Contacting researchers for information on unpublished studies yielded a few abstracts and dissertation monographs. Studies were included if they (1) assessed the effect of sleep deprivation on cognitive function, memory, vigilance, or clinical performance, selected because they are relevant to physi-

**Table 2**—The Effect of Sleep Loss on Performance (Overall and by Subject)

| Analysis             | K   | T     | Subjects, |                                     | Lower<br>95%CI | Upper<br>95%CI | S <sup>2</sup> <sub>1</sub> % |
|----------------------|-----|-------|-----------|-------------------------------------|----------------|----------------|-------------------------------|
|                      |     |       | no.       | $\bar{\delta}$ SE( $\bar{\delta}$ ) |                |                |                               |
| Overall              | 167 | 5,295 | 1,987     | -.951 (.053)                        | -.847          | -1.056         | 72                            |
| MDs                  | 52  | 2,345 | 959       | -.880 (.089)                        | -.705          | -1.054         | 79                            |
| MDs, clin. perf.     | 14  | 635   | 552       | -1.536 (.227)                       | -1.092         | -1.981         | 89                            |
| MDs, excl. cl. perf. | 38  | 1,710 | 883       | -.657 (.103)                        | -.456          | -.858          | 75                            |
| Others               | 115 | 2,950 | 1,028     | -.995 (.064)                        | -.870          | -1.119         | 65                            |

K, number of studies; T, number of effect indexes in sample;  $\bar{\delta}$  and SE( $\bar{\delta}$ ), average effect and standard error corrected for measurement error; CI, confidence interval; S<sup>2</sup><sub>1</sub>, % of remaining variance not accounted for by statistical artifact.

cian functioning;<sup>39,40</sup> (2) involved adults of at least 19 years of age; (3) reported data that could be transformed into a *d* value (the standardized difference between the sleep-deprived and the rested groups); and (4) provided the number of hours participants went without sleep. For studies that assessed the effect of napping, caffeine, and drugs, only the values for the groups that did not receive the intervention and for the non-sleep-deprived controls were included because the goal was to assess the effect of sleep loss in the absence of these alertness-promoting strategies.

The search produced 60 studies that met the selection criteria (listed in Appendix 1).<sup>1,3,26,27,29,30,48,51,58,62-112</sup> All were in English, though several reported research performed in Europe, Asia, and India. Most studies had been published in the scientific literature. In addition to extracting the data to calculate *d* values, each study was coded to record information for the examination of potential moderator variables. Twenty studies used resident physicians as subjects, and 40 involved nonphysician participants. The total sample included 959 physicians and 1,028 nonphysician participants and 5,295 individual effect indexes (2,345 from physician and 2,950 from nonphysicians). Average sample size was 31.2 and ranged from 6 to 424, with 1 large study contributing more than 40% of the overall physician sample. Most studies reported participants' age, and the sample-weighted average was 24.7 years (27.3 for physicians, 22.9 for nonphysicians). Thirty-nine studies analyzed sleep loss of 24 to 30 hours, the period of continuous duty permitted under the ACGME's new duty-hour standards. Chronic partial sleep deprivation was of interest due to its prevalence in residents. Only 6 studies assessed the effects of partial sleep loss, and only 1 of these dealt with resident physicians, although a coexistence of acute and chronic sleep deprivation can be assumed in all studies of postcall residents. Another 61 studies that examined sleep loss and performance had to be excluded from the analysis because the authors did not provide the results that could be transformed into *d* values or the studies lacked control groups or baseline measures. This included several important studies of the effect of acute sleep loss on residents that did not collect data for rested controls.<sup>41,42,52</sup>

### Meta-Analytic Procedures

The National Library of Medicine defines meta-analysis as "a quantitative method of combining the results of independent studies (usually drawn from the published literature) and synthe-

**Table 3**—The Effects of Type of Performance and Extent of Sleep Loss

| Analysis            | T     | Subjects, | $\bar{\delta}$ | Lower<br>95% CI | Upper<br>95% CI | S <sup>2</sup> <sub>1</sub> % |
|---------------------|-------|-----------|----------------|-----------------|-----------------|-------------------------------|
|                     |       |           |                |                 |                 |                               |
| Type of Performance |       |           |                |                 |                 |                               |
| Clin. performance   | 635   | 552       | -1.536         | -1.092          | -1.981          | 89                            |
| Cognitive function  | 2,025 | 1,429     | -.564          | -.406           | -.722           | 67                            |
| Memory              | 1,049 | 815       | -.806          | -.574           | -1.038          | 69                            |
| Vigilance           | 1,610 | 835       | -1.330         | -1.124          | -1.536          | 77                            |
| Sleep Loss, h       |       |           |                |                 |                 |                               |
| 24 to < 30          | 2,899 | 1,218     | -1.142         | -1.006          | -1.278          | 71                            |
| 30 to < 54          | 1,498 | 460       | -.553          | -.332           | -.784           | 78                            |
| 54                  | 337   | 97        | -1.589         | -1.056          | -2.116          | 85                            |
| Chronic             | 561   | 281       | -.886          | -.750           | -1.025          | 24                            |
| Partial             |       |           |                |                 |                 |                               |

T, number of effect indexes;  $\bar{\delta}$ , average effect corrected for measurement error; CI, confidence interval; S<sup>2</sup><sub>1</sub>, percentage of remaining variance not accounted for by statistical artifact.

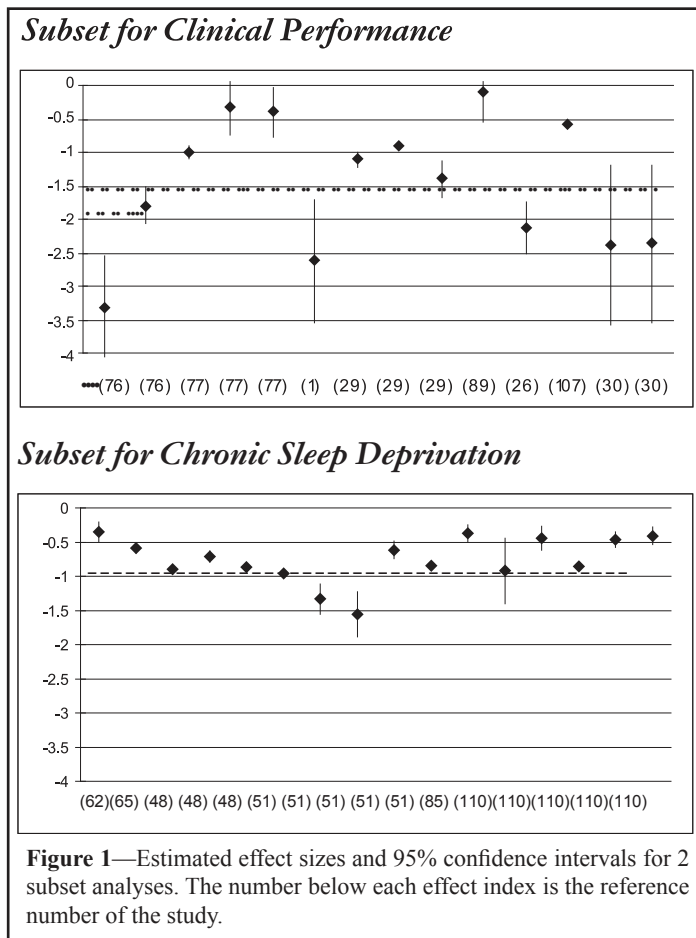
sizing summaries and conclusions.<sup>77,13</sup> This meta-analysis used the Hunter and Schmidt approach, which assesses whether a construct (the effect of sleep loss on cognitive performance) has general validity.<sup>21</sup> It has been used to cumulate the results of studies of human performance in the behavioral and social sciences<sup>114</sup> and is an appropriate method for a meta-analyzing the effects of sleep loss on performance in healthy adults, such as the residents and volunteers in laboratory studies. It was the method used in both earlier meta-analyses of sleep loss and performance.<sup>37,38</sup> Other meta-analytic approaches and statistical models for combining the results of independent studies exist, especially in the clinical trial literature.<sup>115,116</sup> These methods have recently been compared to the Hunter and Schmidt approach.<sup>117</sup>

The measure of effect size used is the *d* value, the difference between the sleep-deprived and the rested group, divided by the pooled standard deviation to allow an unbiased comparison of the results across different studies. *d* Values were computed using the Hunter and Schmidt formula,<sup>21</sup> and their mathematical signs were adjusted as needed to ensure that reduced performance by the sleep-deprived cohort was reflected by a negative value. When studies reported more than 1 effect, all were included, as is commonly done in meta-analysis.<sup>21</sup> Because the focus is on how sleep loss affects the performance of individuals, a repeated-measures design, in which the same individual performs in a sleep-deprived and a rested condition, is the appropriate design for the analysis. Studies using an independent group design were converted using the method described by Morris and DeShon.<sup>118</sup> Studies were weighted by the inverse of their sampling error, giving greater weight to studies with large samples and lower sampling error.<sup>119</sup> The square root of the average reliability of the cognitive tests in the studies was used to adjust the effect indexes for measurement error.<sup>21</sup> Adjustment increased effect indexes by 8% to 14%.

### RESULTS

Sleep Loss and Performance in Physicians and Nonphysicians Table 2 shows the results of the overall meta-analysis and the data for resident physicians and nonphysicians. Across all studies in the analysis, sleep loss reduced cognitive performance by





nearly 1 standard deviation (-.951). When physicians and non-physicians were analyzed separately, the reduction in performance was smaller in physicians (-.880 vs. -.995). Table 3 shows the results disaggregated by type of performance and length and type of sleep deprivation. This shows that the effects of sleep loss for vigilance and clinical performance were larger than those for memory and cognitive function, and the confidence intervals for vigilance and clinical performance and those for memory and cognitive function did not overlap. The analysis distinguished between 3 time increments. Sleep loss of 54 hours or more had the largest effect, but sleep loss of less than 30 hours—the period of continuous duty permitted under the ACGME's standards—had the second-largest effect. Chronic partial sleep loss also resulted in a significant reduction in cognitive performance, with a correct *d* value of -.886.

The confidence intervals in Tables 2 and 3 are an index of the precision of the estimates, which is higher when many studies are combined.<sup>21</sup> A 95% confidence interval reflects a probability of 19 to 1 that the range within the bounds includes the true population value. The last columns in Tables 2 and 3 show the variance after correcting for measurement error. Schmidt and Hunter suggest that remaining variance of 25% or less suggests homogeneity in the effect across studies.<sup>119</sup> The data in Table 3 suggest that unexplained variance remains in all data subsets except for chronic partial sleep loss.

Figure 1 compares data for individual studies and the aggregated effect indexes for studies of clinical performance and chronic sleep deprivation, showing the difference between the 2 data sets, highlighting the greater variance and wider confidence intervals for most studies of clinical performance, compared to studies of

**Table 4**—Hierarchical Analysis by Subject, Type of Performance and Degree of Sleep Loss

| Analysis                | 24 to < 30 hours Acute Sleep Loss |          |           | Chronic Partial Sleep Loss |     |          |           |       |
|-------------------------|-----------------------------------|----------|-----------|----------------------------|-----|----------|-----------|-------|
|                         | T                                 | $\delta$ | $S^2_1$ % | Q                          | T   | $\delta$ | $S^2_1$ % | Q     |
| MIDs                    | 1,862                             | -.986    | 77        | 178.9                      | --  | --       | --        | --    |
| Excl. clin. performance | 1,227                             | -.701    | 59        | 66.3                       | --  | --       | --        | --    |
| Cognitive function      | 697                               | -.560    | 81        | 42.6                       | --  | --       | --        | --    |
| Memory                  | 316                               | -.771    | 21        | 15.2                       | --  | --       | --        | --    |
| Vigilance               | 214                               | -.904    | 63        | 18.9                       | --  | --       | --        | --    |
| Clin. Performance       | 635                               | -1.536   | 89        | 128.9                      | --  | --       | --        | --    |
| Non-physicians          | 1,037                             | -1.331   | 66        | 189.3                      | 294 | -.886    | 24        | 12.9* |
| Cognitive function.     | 509                               | -.704    | 20        | 34.8*                      | --  | --       | --        | --    |
| Memory                  | 178                               | -1.237   | 51        | 26.5                       | 46  | -.794    | 03        | 3.1*  |
| Vigilance               | 350                               | -2.083   | 86        | 174.9                      | 181 | -.977    | 14        | 7.0*  |

T, number of effect indexes in sample;  $\delta$ , average effect corrected for measurement error;  $S^2_1$ , % of remaining variance not accounted for by statistical artifact; Q, Q statistic for homogeneity in true effect across studies.

\*Indicates Q statistic not significant at the .005 level, suggesting homogeneity.

chronic partial sleep deprivation.

The finding of a -.886 reduction in performance for chronic sleep loss is consistent with the literature.<sup>46-51</sup> It contrasts with Pilcher and Huffcutt's finding of a -3.01 standard deviation reduction in performance.<sup>38</sup> A partial explanation is that 4 of 6 studies in their analysis were of postcall residents, which this meta-analysis assigned to the rubric short-term sleep loss, due to the assumed underlying chronic sleep deprivation in the "rested" controls under in these studies. Alternative assignment of these studies to the "chronic partial sleep loss" category increases the negative effect somewhat (-.951, data not shown), but it remained smaller than that found by Pilcher and Huffcutt.

Legend for Figure 1: Shows estimated effect indexes (corrected *d* values) and 95% confidence intervals for the outcomes of individual studies included in the subset and subset pooled effect size (shown by dotted line).

### Results of the Hierarchical Analysis of Moderators

Table 4 shows the results for the comprehensive moderator analysis for the 2 types of sleep loss most relevant to resident performance under the current ACGME standards. It shows data for physicians versus nonphysicians grouped by type of performance under sleep loss of 24 to 30 hours and chronic sleep deprivation, showing the combined effect of several moderators. The effect of sleep loss of 24 to 30 hours was smaller in physicians than in nonphysicians, and, when clinical performance was excluded, the difference became statistically significant (.05 level). Clinical performance showed the largest decrement in physicians (-1.536). Vigilance showed the largest decline in nonphysicians (-2.083). Chronic partial sleep loss in physicians could not be meta-analyzed, as this group included only 1 study.

Table 4 shows 2 indexes for the homogeneity of the effect across studies, the percentage of variance not attributed for by sampling and measurement error ( $S^2_1$  %) and a Q statistic.<sup>21</sup> For 4

**Table 5**—“Study Factors” for Studies Including Less Than 30 Hours of Sleep Loss in Physicians

| Authors, Year                          | Control | Definition of Sleep Loss/Rested   | Quality of Assessment  |
|--|---------|---|--|
| Browne et al, 1994 <sup>27</sup>       | Low     | Postcall; less than 4 h of sleep, rested greater than 4 h of sleep.   | Participants kept sleep logs; test used articles to assess short-term recall (practice effect not an issue).                                       |
| Deaconson et al, 1988 <sup>70</sup>    | Low     | Postcall; less than 4 h of sleep, rested greater than 4 h of sleep.   | Participants kept sleep logs and recorded subjective fatigue.  |
| Eastridge et al, 2003 <sup>76</sup>    | Low     | Postcall; no difference in mean h of sleep between groups ( $P = .17$ ).  | Participants tested between 8 and 11 am; participants kept sleep logs; recorded subjective fatigue.  |
| Engel et al, 1987 <sup>77</sup>        | High    | Postcall; mean 3.6 h sleep, rested averaged 7.3 h sleep.  | Testing involved simulated patient history. No added recording of conditions at testing.   |
| Fluck et al, 1998 <sup>78</sup>        | High    | Postcall; mean sleep of $2.75 \text{ h} \pm 38 \text{ min}$ over 28 h, rested not defined.                            | Subjects' sleepiness was rated in the rested and sleep-deprived condition.   |
| Friedman et al, 1971 <sup>1</sup>      | High    | Postcall; mean 1.8 h sleep in study group, rested averaged 7 h sleep.   | Time of test administration varied; incentive offered to interns who made fewest mistakes.   |
| Grantcharov et al, 2001 <sup>29</sup>  | High    | Post “night-shift;” less than 3 h of sleep, rested not defined.   | All participants tested at 9:30 am, testing scheme thought to reduce practice effects.   |
| Hart et al, 1987 <sup>84</sup>         | High    | Postcall; mean 2.7 h sleep, rested average 7.9 h sleep.   | Resident recorded the number of hours slept for 2 nights prior to testing.   |
| Hawkins et al, 1985 <sup>85</sup>      | Low     | Postcall; less than 5 h of sleep, rested greater than 5 h of sleep.   | Residents were dually categorized by their acute and chronic sleep deprivation.  |
| Howard et al, 2003 <sup>89</sup>       | High    | Participants kept aware at least 25 h prior to testing on simulated case; rested relieved of duties until 10 am.      | Four-hour simulated anesthesia case; sleep monitored via wrist activity monitors and sleep log; sleepiness assessed via Stanford Sleepiness Scale. |
| Jacques et al, 1990 <sup>90</sup>      | Low     | Postcall; sleep deprived and rested not further defined.  | Tested using “in-training exam, practice effect may not be an issue. No assessment of hours of sleep.  |
| Kannan et al, 1997 <sup>92</sup>       | Low     | Postcall; Less than 4 h of sleep, rested not defined.   | Testing spaced 1 week apart to reduce practice effect.   |
| Lingenfelter et al, 1994 <sup>26</sup> | High    | Postcall; sleep-deprived not further defined, rested more than 6 h of sleep.  | Residents kept sleep diary; all participants tested at 8:00 am.  |
| Mak and Spurgeon, 2003 <sup>96</sup>   | Low     | Postcall, slept mean 2.9 (1.0) h during 29.3 (3.8) h of call, sleep prior to testing in rested condition not defined. | Testing in sleep-deprived and rested condition spaced out a mean of 7 days.  |
| Reznick and Folse, 1987 <sup>103</sup> | Low     | Postcall; less than 3 h of sleep, rested greater than 5 h of sleep.   | Participants tested between 8 am and 6 pm, only 12 of 21 participants completed testing in both conditions.  |
| Richardson et al, 1996 <sup>104</sup>  | High    | Postcall; tested on the morning after call; study group average 3.7 h of sleep.                                       | Sleep-deprived tested in the morning of post-call day; sleep recorded via electroencephalogram.  |
| Stone et al, 2000 <sup>107</sup>       | Low     | Postcall; sleep-deprived and rested not further defined.  | Tested using ABSITE exam, practice effect may not be an issue. No assessment of hours of sleep.  |
| Taffinder et al, 1998 <sup>30</sup>    | High    | Sleep-deprived tested after night of no sleep; rested after an “undisturbed night.”                                   | Participants were pre-trained in the simulator to reduce practice effects, further controlled by Latin square design.                              |

ABSITE refers to the American Board of Surgery In-Training Exam.

subsets—cognitive performance in nonphysicians under less than 30 hours of sleep loss and cognitive performance, memory, and vigilance in nonphysicians under partial chronic sleep deprivation—both indexes suggested homogeneity. Unexplained variance remains in all other sets, suggesting residual study factors or moderators. Figure 2 presents the data for short-term sleep deprivation, disaggregated by subject and type of performance, highlighting the similarity in the patterns for physicians and non-physicians and the smaller effect sizes in studies of physicians.

### The Role of “Study Factors”

Some studies of physicians found little or no reduction in per-

formance for the sleep-deprived group. One objective of this meta-analysis is to explore whether this was due to “study factors,” as suggested in a qualitative review in 2002.<sup>40</sup> Nearly all studies in nonphysicians were conducted in sleep laboratories. In contrast, most data on sleep loss in physicians came from field studies in which postcall residents were compared with their colleagues. To assess the implications of the different approaches, the analysis sought information on whether field studies of residents included in the meta-analysis used controls that were similar to laboratory studies. Examples encompass exact recording of the number of hours subjects went without sleep, assessing sleepiness prior to testing, and allowing pretest practice to guard against practice effects that may mask reduced performance in sleep-deprived subjects.

Table 5 categorizes the 15 studies of sleep deprivation of less than 30 hours duration in physicians that quantified the hours of sleep or wakefulness into “high control” and “low control” studies. Another 3 physician studies did not offer information on the hours the “sleep-deprived” postcall group had gone without sleep. Of the 15 studies, those that offered clear definitions of “sleep-deprived” and “rested” and assessed sleepiness or used other study controls were designated as “high” control, the remainder was designated as “low” control. Figure 3 shows that, for the 11 effect indexes for cognitive function and memory, the sample-weighted average effect size was larger in high control studies (-1.16 vs -.25). This suggests a relationship between the use of controls that make field studies more like studies conducted in sleep laboratories and effect size. Only studies assessing cognitive function and memory were included to provide a homogenous set of effects for this analysis.

Added support for the role of study factors comes from the results of an alternative analysis of the data for physicians that excluded the findings of 1 study with a very large sample (424 participants)<sup>107</sup> classified as a “low” control study. When data for this study were excluded, the performance reduction in sleep-deprived physicians became comparable with that of nonphysicians (corrected  $d$  value = -.965, compared to -.995), and the decrement for clinical performance increased to -2.084 (data not shown). Alternative calculations that exclude data from studies with very large sample sizes are commonly done in meta-analysis to assess whether the results of 1 study bias the findings.

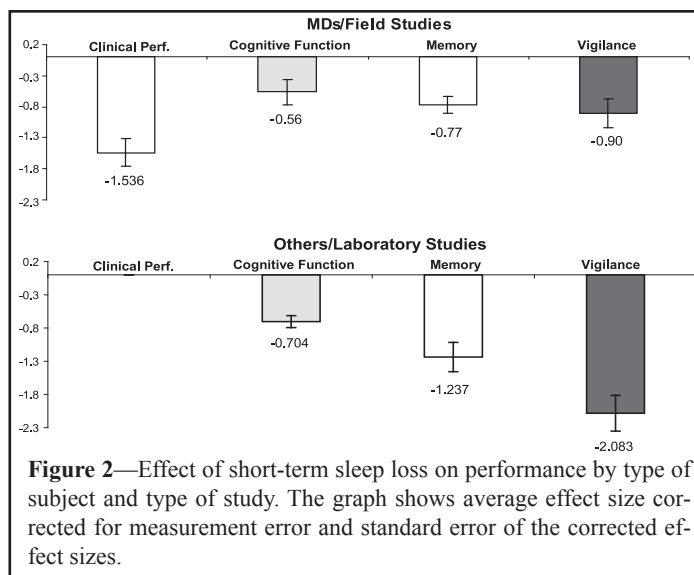
## DISCUSSION

This meta-analysis found that cognitive performance in physicians is affected by sleep deprivation, confirming the results of earlier literature reviews.<sup>39,40</sup> Smaller effect sizes in physician studies likely relate to difficulty in field studies to control the exact hours of sleep loss or to compensate for the chronic reduced sleep in the “rested” cohort, not differences in how physicians and nonphysicians respond to sleep loss. In many studies of medical residents included in this analysis, the rested controls met the clinical definition for chronic sleep deprivation (less than 5 to 6 hours of sleep for a period of several nights), and both acute and chronic sleep loss likely were present in the study group.<sup>40,45,120</sup>

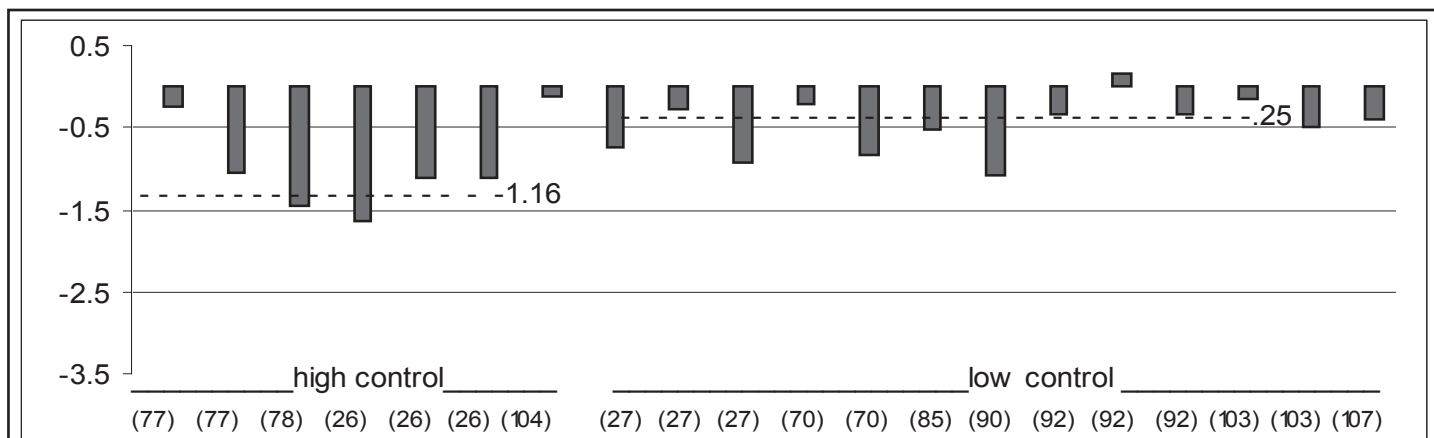
The analysis showed that components of performance important to the work of physicians respond differently to sleep deprivation. The published literature has suggested that vigilance is

most susceptible,<sup>121-124</sup> but this analysis found a negative effect on clinical performance that was 70% greater than that for vigilance. This finding is important. It represents the first aggregation of studies of the effect of sleep deprivation on physicians’ clinical performance and counters commentary that residents’ reduced performance on cognitive tests would not translate into reduced clinical performance.<sup>125</sup> At the same time, the results are tentative because of the limited sample and the heterogeneity of studies, which included clinical tasks, tests of clinical performance, and “high-fidelity” clinical simulations.

The findings of the effect of sleep loss on residents’ cognitive and clinical performance have implications for the effort to set limits on resident hours. Both acute and chronic partial sleep loss is prevalent in many residents, and the ACGME minimum duty-hour standards seek to address both. Chronic sleep loss has been formally addressed in just 1 study of physicians, which found reduced performance only for acute sleep loss.<sup>85</sup> Another focal area is the length of the call and the postcall period, given the continued emphasis on in-house call in many specialties. The ACGME limits continuous duty to 24 hours plus a period of up to 6 hours for transfer of care and didactics. The finding that sleep loss of 24 to 30 hours produces a -.986 reduction in physicians’ aggregate cognitive and clinical performance suggests that residents may experience significant performance decrements under the current minimum standards. Applying Picher and Huffcutt’s pragmatic



**Figure 2**—Effect of short-term sleep loss on performance by type of subject and type of study. The graph shows average effect size corrected for measurement error and standard error of the corrected effect sizes.



**Figure 3**—Estimated effect sizes and 95% confidence intervals for 2 subset analyses. The number below each effect index is the reference number of the study.

interpretation, individuals at the middle of the cohort with 24 hours to 30 hours of sleep deprivation performed comparably to individuals at the 15<sup>th</sup> percentile of the rested group.<sup>38</sup> For clinical performance ( $d = -1.536$ ), sleep-deprived physicians performed at the seventh percentile of the comparison group. Since the ACGME established its limit on continuous duty in 2003, there have been requests to relax it in the interest of maintaining continuity of care, acculturating residents to the expectations of independent practice, and preserving time for clinical learning.<sup>126-128</sup> There also have been calls to further reduce the maximum continuous hours permitted, with support provided by the findings of 2 studies in which a 16-hour limit increased alertness and reduced medication errors by first-year residents rotating in a critical-care setting.<sup>41,42</sup>

## Limitations

This study has limitations common to meta-analysis, including the inability to include all studies that investigate a given phenomenon, because of data, technical, and other exclusion factors. Another limitation may be the small number of nonpublished studies included. While there is disagreement about the importance of unpublished data, some feel that excluding it results in a bias toward studies that produce significant findings in the expected direction.<sup>129</sup> The author's conscientious search for nonpublished materials and the inclusion of nonsignificant and positive effect indexes should reduce this concern. The analysis solely is concerned with the number of hours without sleep, although the effect of sleep loss on performance is more complex. It is influenced by time of day and circadian rhythm, which produce peaks and valleys in performance throughout the day that this analysis was not able to assess.<sup>69,86,130,131</sup> Four aspects of sleep deprivation of interest to performance in physicians—chronic partial sleep deprivation and the effects of task duration, task pacing, and task complexity—could not be analyzed due to a lack of primary studies. Finally, this meta-analysis offers no evidence about the prevalence of sleep loss in residents. Studies done prior to the ACGME's institution of duty-hour limits found that many residents averaged more than 100 hours per week of work.<sup>132,133</sup> Data under the new limits are being collected and suggest that hours have declined in specialties that previously exceeded 80 weekly hours.<sup>134</sup>

## Areas for Future Research

Further research is needed in several areas to enhance our understanding of the effects of sleep loss in residents. One area is chronic partial sleep deprivation, common in residency, yet only 1 study assessed its effect in physicians.<sup>85</sup> Another involves the hours of continuous service, an area in which studies are beginning to raise concerns with the current length of the on-call and postcall period.<sup>41,42,44</sup> Given the presumed continued prevalence of chronically reduced sleep and long continuous duty periods in many residencies, research is needed to find the limit on weekly and continuous hours and the organization of call and continuity experiences that will optimally foster the dual goals of resident learning and safe and effective patient care. Research also should analyze the effect of sleep deprivation on learning, a primary goal of residency.

No meta-analysis of sleep loss in physicians would be complete without discussing the link to human error. In 1 study, 41% of residents reported that their sleep deprivation had contributed

to a serious medical error.<sup>135</sup> The reductions in clinical performance found in this meta-analysis highlight the need for further study of the clinical consequences of sleep loss and the benefits of duty-hour limits to patient safety. This research will not be easy to conduct or interpret. An analysis of the effect of interventions to reduce resident hours concluded that the evidence was insufficient to determine that these interventions improved patient safety,<sup>136</sup> and a retrospective review of New York's statewide database to assess the benefits of its state-imposed limits found that reductions in mortality were comparable in teaching and nonteaching institutions and could not be attributed to the limits on resident hours.<sup>137</sup> Errors result from a combination of individual and system factors,<sup>138,139</sup> and resident performance must be viewed in the context of a host of factors in their learning environment that may enhance or detract from patient safety. Duty-hour limits are 1 important aspect. Research is needed to enhance the educational community's ability to adapt patient care and learning models to the reductions in resident hours. This includes efforts to improve the systems for patient hand-off and to counteract the effects of reduced practice in surgical residencies and diminished continuity of care.<sup>140,141</sup> Finally, there is a need to find the best ways to educate residents and teaching faculty about the science of sleep and performance, recognition of sleep deprivation and fatigue, and strategies to promote alertness.<sup>11,142</sup>

## ACKNOWLEDGMENTS

The author would like to thank Dr. Frank L. Schmidt for his helpful suggestions on drafts of the manuscript.

## REFERENCES

1. Friedman RC, Bigger JT, Kornfield DS. The intern and sleep loss. *N Engl J Med* 1971;285:201-3.
2. Dawson D, Reid K. Fatigue, alcohol and performance impairment. *Nature* 1997;88:235.
3. Lamond N, Dawson D. Quantifying the performance impairment associated with fatigue. *J Sleep Res* 1999;8:255-262.
4. Public Citizen. Petition to the Occupational Safety and Health Administration (OSHA), filed by Public Citizen, the American Medical Student Association and the Committee of Interns and Residents (CIR); Washington, DC; April 30, 2001.
5. 107th CONGRESS, 1st Session, H. R. 3236. The Patient and Physician Safety and Protection Act of 2001; Washington, DC, November 6, 2001.
6. Geer RT, Jobes DR, Tew JD, Stepsis LH. Incidence of automobile accidents involving anesthesia residents after on-call duty cycles. *Anesthesiology* 1997;87:A938.
7. Marcus CL, Loughlin GM. Effect of sleep deprivation on driving safety in housestaff. *Sleep* 1996;19:763-6.
8. Steele MT, Ma OJ, Watson WA, Thomas HA Jr, Muelleman RL. The occupational risk of motor vehicle collisions for emergency medicine residents. *Acad Emerg Med* 1999;6:1050-3.
9. Parks DK, Yetman RJ, McNeese MC, Burau K, Smolensky MH. Day-night pattern in accidental exposures to blood-borne pathogens among medical students and residents. *Chronobiol Int*. 2000;17:61-70.
10. Accreditation Council for Graduate Medical Education, Common program requirements for duty hours. Approved June 24, 2003. Available at: <http://www.acgme.org/DutyHours/dutyHoursCommonPR.pdf>. Accessed August 3, 2005.
11. Accreditation Council for Graduate Medical Education. Report of the ACGME Work Group on Resident Duty Hours, page 5. June 11, 2002. Available at: <http://www.acgme.org/DutyHours/>



- wkgroupreport611.pdf. Accessed August 3, 2005.
12. Association of American Medical Colleges. AAMC policy guidance on graduate medical education, September 2001. Available at: <http://www.aamc.org/patientcare/gmepolicy/start.htm>. Accessed August 3, 2005.
  13. Asch DA, Parker RM. The Libby Zion case. One step forward or two steps backward? *N Engl J Med* 1988;318:771-5.
  14. Evans L. Regulatory and legislative attempts at limiting medical resident work hours. *J Leg Med* 2002;23:251-67.
  15. Naitoh P. Sleep deprivation in human subjects: a reappraisal. *Waking Sleeping* 1976;1:53-60.
  16. Leung L, Becker CE. Sleep-deprivation and house staff performance—Update 1984-1991. *J Occup Med* 1992;34:1153-60.
  17. Samkoff JS, Jacques CHM. A review of studies concerning effects of sleep-deprivation and fatigue on residents' performance. *Acad Med* 1991;66:687-93.
  18. Schmidt FL, Hunter JE, Urry VE. Statistical power in criterion-related validation studies. *J Appl Psychol* 1976;61:473-85.
  19. Sedlmeier P, Gigerenzer G. Do studies of statistical power have an effect on the power of the studies? *Psych Bull* 1985;105:309-16.
  20. Hedges LV. How hard is hard science, how soft is soft science: the empirical cumulativeness of research. *Am Psychol* 1987;42:443-55.
  21. Hunter, JE, Schmidt FL. *Methods of Meta-analysis*. Newbury Park, CA: Sage Publications; 1990.
  22. Schwartz AJ, Black ER, Goldstein MG, et al: Levels and causes of stress among residents. *J Med Educ* 1987;62:744-753.
  23. Asken MJ, Raham DC: Resident performance and sleep deprivation: a review. *J Med Educ* 1983;58:382-388.
  24. Garcia EE: Sleep deprivation in physician training. *NY State J Med* 1987;87:637-638.
  25. Page L. Long work weeks typical for residents--survey. *Am Med News*, February 5, 1988, pp 10-12.
  26. Lingenfelter T, Kaschel R, Weber A, Zaiser-Kaschel H, Jakober B, Kuper J. Young hospital doctors after night duty – their task-specific cognitive status and emotional condition. *Med Educ*. 1994;28:566-572.
  27. Browne BJ, VanSusteren T, Onsager DR, Simpson D, Salaymeh B, Condon RE. Influence of sleep-deprivation on learning among surgical house staff and medical students. *Surgery*. 1994;115:604-610.
  28. Haynes DF, Schwedler M, Dyslin DC, Rice JC, Kerstein MD. Are postoperative complications related to resident sleep deprivation? *South Med J*. 1995;88:283-289
  29. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Laparoscopic performance after one night on call in a surgical department: prospective study. *BMJ*. 2001;323:1222-3.
  30. Taffinder NJ, McManus IC, Gul Y, et al. Effect of sleep deprivation on surgeons' dexterity on laparoscopy simulator. *Lancet*. 1998;352:1191.
  31. Aya AGM, Mangin R, Robert C, Ferrer JM, Eledjam JJ. Increased risk of unintentional dural puncture in night-time obstetric epidural anesthesia. *Can J Anaesth*. 1999;46:665-69.
  32. Storer JS, Floyd HH, Gill WL, et al. Effects of sleep deprivation on cognitive ability and skills of paediatric residents. *Acad Med* 1989;64:29-32.
  33. Ben-Aroya Z, Segal D, Hadar A, et al. Effect of Ob/Gyn residents' fatigue and training level on the accuracy of fetal weight estimation. *Fetal Diagn Ther* 2002;177-81.
  34. Smith-Coggins R, Rosekind MR, Hurd S, Buccino KR. Relationship of day versus night sleep to physician performance and mood. *Ann Emerg Med* 1994;24:928-34.
  35. Denisco RA, Drummond JN, Gravenstein JS. The effect of fatigue on the performance of a simulated anesthetic monitoring task. *J Clin Monit* 1987;3:22-4.
  36. Bertram DA. Characteristics of shifts and second-year resident performance in an emergency department. *NY State J Med* 1988;88:10-4.
  37. Koslowsky M, Babkoff H. Meta-analysis of the relationship between total sleep deprivation and performance. *Chronobiol Int* 1992;9:132-6.
  38. Pilcher JJ, Huffcutt AI. Effects of sleep deprivation on performance: A meta-analysis. *Sleep* 1996;19:318-26.
  39. Weinger MB, Ancoli-Israel S. Sleep deprivation and clinical performance. *JAMA* 2002;287:955-57.
  40. Veasey S, Rosen R, Barzansky B, Rosen I, Owens J. Sleep loss and fatigue in residency training: a reappraisal. *JAMA* 2002;288:1116-24.
  41. Lockley SW, Cronin JW, Evans EE, et al. Harvard Work Hours, Health and Safety Group. Effect of reducing interns' weekly work hours on sleep and attentional failures. *N Engl J Med* 2004;351:1829-37.
  42. 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.
  43. Gottlieb DJ, Parenti CM, Peterson CA, Lofgren RP. Effect of a change in house staff work schedule on resource utilization and patient care. *Arch Intern Med* 1991;151:2065-70.
  44. Van Dongen HPA, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep* 2003;26:117-26.
  45. Carskadon M, Dement WC. Cumulative effects of sleep restriction on daytime sleepiness. *Psychophysiology* 1981;18:107-13.
  46. Baldwin DC Jr, Daugherty SR. Sleep Deprivation and fatigue in residency training: results of a national survey of first- and second-year residents. *Sleep* 2004;27:217-23.
  47. Jewett ME, Dijk DJ, Kronauer RE, Dinges DF. Dose-response relationship between sleep duration and human psychomotor vigilance and subjective alertness. *Sleep* 1999;22:171-9.
  48. Blagrove AU, Blagrove, M, Alexander, C, Horne, JA. The effects of chronic sleep reduction on the performance of cognitive tasks sensitive to sleep-deprivation. *Appl Cogn Psychol* 1995;9:21-40.
  49. Dinges D, Maislin G, Van Dongen H. Chronic sleep restriction: relation of sleep structure to daytime sleepiness and performance. *Sleep* 2001;24:A28.
  50. Webb WB, Agnew HW. The effects of a chronic limitation of sleep length. *Psychophysiology* 1974;11:265-74.
  51. Dinges DF, Pack F, Williams K et al. Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep* 1997;20:267-77.
  52. Papp KK, Stoller EP, Sage P, et al. The effects of sleep loss and fatigue on resident physicians: a multi- institutional, mixed-method survey. *Acad Med* 2004;79:394-406.
  53. Howard SK, Gaba DM, Rosekind MR, Zarcone VP. The risks and implications of excessive daytime sleepiness in resident physicians. *Acad Med* 2002;77:1019-25.
  54. Bird CM, Papadopoulou K, Ricciardelli P, Rossor MN, Cipolotti L. Test-retest reliability, practice effects and reliable change indices for the recognition memory test. *Br J Clin Psychol* 2003;42:407-25.
  55. Rawlings DB, Crewe NM. Test-retest practice effects and test score changes of the WAIS-R in recovering traumatically brain-injured survivors. *Clin Neuropsychol* 1992;6:415-30.
  56. Connor A, Franzen M, Sharp B. Effects of practice and differential instructions on Stroop performance. *Int J Clin Neuropsychol* 1988;10:1-4.
  57. 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:A485.
  58. Caldwell JA, Ramspott S. Effects of task duration on sensitivity to sleep deprivation using the multi-attribute task battery. *Behav Res Meth Instrum Comp* 1998;30:651-60.



59. Van Dongen HP, Baynard MD, Nosker GS, Dinges DF. Repeated exposure to total sleep deprivation: Substantial trait differences in performance impairment among subjects. *Sleep* 2002;25:121-123.
60. Aeschbach D, Sher L, Postolache TT, Matthews JR, Jackson MA, Wehr TA. A longer biological night in long sleepers than in short sleepers. *J Clin Endocrinol Metab* 2003;88:26-30.
61. Van Dongen HP, Maislin G, Dinges DF. Dealing with inter-individual differences in the temporal dynamics of fatigue and performance: importance and techniques. *Aviat Space Environ Med* 2004;75:A147-54.
62. Anderson CM, Maislin G, Van Dongen HP, et al. Effect of chronically reduced nocturnal sleep, with and without daytime naps, on neurobehavioral performance. *Sleep* 2000;23:A74.
63. Angus RG, Heslegrave RJ. Effects of sleep loss on sustained cognitive performance during a command and control simulation. *Behav Res Meth Instrum Comput* 1985;17:55-67.
64. Beaumont M, Batejat D, Pierard C, et al. Slow-release caffeine and prolonged (64-h) continuous wakefulness: effects on vigilance and cognitive performance. *J Sleep Res* 2001;10:265-276.
65. Belenky G, Wesensten NJ, Thorne DR, et al. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *J Sleep Res* 2003;12:1-12.
66. Bonnet MH. The effect of varying prophylactic naps on performance, alertness and mood throughout a 52-hour continuous operation. *Sleep* 1991;14:307-15.
67. Bonnet MH, Gomez S, Wirth O, Arand DL. The use of caffeine vs prophylactic naps in sustained performance. *Sleep* 1995;18:97-104.
68. Caldwell JA, LeDuc PA. Gender influences on performance, mood and recovery sleep in fatigued aviators. *Ergonomics* 1998;41:1757-71.
69. Casagrande M, Violani C, Curcio G, Bertini M. Assessing vigilance through a brief pencil and paper letter cancellation task (LCT): Effects of one night of sleep deprivation and of the time of day. *Ergonomics* 1997;40:613-30.
70. Deaconson TF, O'Hair DP, Levy MF, Lee MBF, Schueneman AL, Condon RE. Sleep-deprivation and resident performance. *JAMA* 1988;260:1721-7.
71. DeGennaro L, Ferrara M, Curcio G, Bertini M. Visual search performance across 40 h of continuous wakefulness: Measures of speed and accuracy and relation with oculomotor performance. *Physiol Behav* 2001;74:197-204.
72. Doran SM, Van Dongen HPA, Dinges DF. Sustained attention performance during sleep deprivation: Evidence of state instability. *Arch Ital Biol* 2001;139:253-67.
73. Dorrian J, Lamond N, Dawson D. The ability to self-monitor performance when fatigued. *J Sleep Res* 2000;9:137-44.
74. Drummond SPA, Brown GG, Stricker JL, Buxton, RB, Wong, EC, Gillin JC. Sleep deprivation-induced reduction in cortical functional response to serial subtraction. *Neuroreport* 1999;10:3745-8.
75. Drummond SPA, Gillin JC, Brown GG. Increased cerebral response during a divided attention task following sleep deprivation. *J Sleep Res* 2001;10:85-92.
76. Eastridge BJ, Hamilton EC, O'Keefe GE, et al. Effect of sleep deprivation on the performance of simulated laparoscopic surgical skill. *Am J Surg* 2003;186:169-74.
77. Engel W, Seine R, Powell V, D'Alessandri R. Clinical performance of interns after being on-call. *South Med J* 1987;80:761-763.
78. Fluck E, File SE, Springett J, et al. Does the sedation resulting from sleep deprivation and lorazepam cause similar cognitive deficits? *Pharmacol Biochem Behav* 1998;59:909-15.
79. Forest G, Godbout R. Effects of sleep deprivation on performance and EEG spectral analysis in young adults. *Brain Cogn* 2000;43:195-200.
80. Gillberg M, Kecklund G, Axelsson J, Akerstadt T. Counteracting sleepiness with a short nap. *J Sleep Res* 1994;3:90.
81. Gottlieb DJ, Peterson CA, Parenti CM, Lofgren RP. Effect of a night float system on housestaff neuropsychological function. *J Gen Intern Med* 1993;8:146-8.
82. Harrison Y, Horne JA. One night of sleep loss impairs innovative thinking and flexible decision making. *Organ Behav Hum Decis Process* 1999;78:128-45.
83. Harrison Y, Horne JA. Sleep loss and temporal memory. *Q J Exp Psychol A* 2000;53:271-9.
84. Hart RP, Buchsbaum DG, Wade JB, Hamer RM, Kwentus JA. Effect of sleep deprivation on first-year residents' response times, memory, and mood. *J Med Educ* 1987;52:940-2.
85. Hawkins MR, Vichick DA, Silsby HD, Kruzich DJ, Butler R. Sleep deprivation and performance of house officers. *J Med Educ* 1985;60:530-5.
86. Heuer H, Spijkers W, Kiesswetter E, Schmidtke V. Effects of sleep loss, time of day, and extended mental work on implicit and explicit learning of sequences. *J Experiment Psychol Appl* 1998;4:139-62.
87. Hockey GRJ, Wastell DG, Sauer J. Effects of sleep deprivation and user interface on complex performance: a multilevel analysis of compensatory control. *Hum Fact* 1998;40:233-53.
88. Horne JA. Sleep loss and "divergent thinking" ability. *Sleep* 1988;11:528-36.
89. Howard SK, Gaba DM, Smith BE, et al. Simulation study of rested versus sleep-deprived anesthesiologists. *Anesthesiology* 2003;98:1345-55.
90. Jacques CH, Lynch JC, Samkoff JS. The effects of sleep loss on cognitive performance of resident physicians. *J Fam Pract* 1990;30:223-9.
91. Jaskowski P, Wlodarczyk D. Effect of sleep deficit, knowledge of results, and stimulus quality on reaction time and response force. *Percept Mot Skills* 1997;84:563-72.
92. Kannan S, Malhotra SK, Bajaj A, Pershad D, Chari P. Sleep deprivation in anaesthetists: Effect on cognitive functions. *Neurol India* 1997;45:250-2.
93. Kim DJ, Lee HP, Kim MS, et al. The effect of total sleep deprivation on cognitive functions in normal adult male subjects. *Int J Neurosci* 2001;109:127-37.
94. Linde L, Edland A, Bergstrom M. Auditory attention and multiattribute decision-making during a 33 h sleep-deprivation period: mean performance and between-subject dispersions. *Ergonomics* 1999;42:696-713.
95. McCarthy ME, Waters WF. Decreased attentional responsivity during sleep deprivation: Orienting response latency, amplitude, and habituation. *Sleep* 1997;20:115-23.
96. Mak SK, Spurgeon P. The effects of acute sleep deprivation on performance of medical residents in a regional hospital: prospective study. *Hong Kong Med J* 2004;10:14-20.
97. Manly T, Lewis GH, Robertson IH, Watson PC, Datta AK. Coffee in the cornflakes: time-of-day as a modulator of executive response control. *Neuropsychologia* 2002;40:1-6.
98. Matsumoto Y, Mishima K, Satoh K, Shimizu T, Hishikawa Y. Physical activity increases the dissociation between subjective sleepiness and objective performance levels during extended wakefulness in human. *Neurosci Lett* 2002;326:133-6.
99. Neri DF, Wiegmann D, Shappell SA, McCardie A, McKay DL. The effects of tyrosine on cognitive performance during extended wakefulness. *Aviat Space Environ Med* 1995;66:313-9.
100. Pigeau R, Naitoh P, Buguet A, McCann C. Modafinil, d-amphetamine and placebo during 64 hours of sustained mental work. Effects on mood, fatigue, cognitive performance and body temperature. *J Sleep Res* 1995;4:212-28.
101. Pilcher JJ, Walters AS. How sleep deprivation affects psychological variables related to college students' cognitive performance. *J Am Coll Health* 1997;45:121-6.
102. Polzella DJ. Effects of sleep deprivation on short-term recognition memory. *J Exp Psychol Hum Learn Mem* 1975;104:194-200.
103. Reznick RK, Folse JR. The effects of sleep-deprivation on performance of surgical residents. *Am J Surg* 1987;154:520-5.
104. Richardson GS, Wyatt JK, Sullivan JP, et al. Objective assessment

- of sleep and alertness in medical house staff and the impact of protected time for sleep. *Sleep* 1996;19:718-26.
105. Robbins J, Gottlieb F. Sleep deprivation and cognitive testing in internal medicine house staff. *West J Med* 1990;12:82-86.
  106. Steyvers FJ. The influence of sleep deprivation and knowledge of results on perceptual encoding. *Acta Psychol* 1987;66:173-87.
  107. Stone MD, Doyle J, Bosch RJ, Bothe A, Steele G. Effect of resident call status on ABSITE performance. *Surgery* 2000;128:465-71.
  108. Thomas M, Sing H, Belenky G et al. Neural basis of alertness and cognitive performance impairments during sleepiness. I. Effects of 24 h of sleep deprivation on waking human regional brain activity. *J Sleep Res* 2000;9:335-52.
  109. Webb WB. A further analysis of age and sleep deprivation effects. *Psychophysiology* 1985;22:156-61.
  110. Wesensten NJ, Belenky G, Kautz MA, Thorne DR, Reichardt RM, Balkin TJ. Maintaining alertness and performance during sleep deprivation: modafinil versus caffeine. *Psychopharmacology* 2002;159:238-47.
  111. Wimmer F, Hoffmann RF, Bonato RA, Moffitt AR. The effects of sleep-deprivation on divergent thinking and attention processes. *J Sleep Res* 1992;1:223-30.
  112. Wright KP, Badia P, Myers BL, Plenzler SC. Combination of bright light and caffeine as a countermeasure for impaired alertness and performance during extended sleep deprivation. *J Sleep Res* 1997;6:526-35.
  113. Selden CR. *Current Bibliographies in Medicine* 92-13: Meta-Analysis; January 1980 through December 1992. Bethesda: National Library of Medicine; 1992.
  114. Schmidt FL, Hunter J. The validity and utility of selection methods in personnel psychology. Practical and theoretical implications of 85 years of research findings. *Psychol Bull* 1998;124:262-74.
  115. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clin Trials* 1986;7:177-88.
  116. Hedges LV, Vevea JL. Fixed and random-effects models in meta-analysis. *Psychological Methods*, 1998;3:486-77.
  117. Hall SM, Brannick MT. Comparison of two random-effects methods of meta-analysis. *J Appl Psychol* 2002;87:377-89.
  118. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-group designs. *Psych Methods* 2002;7:105-25.
  119. Schmidt FL, Hunter J. Development of a general solution to the problem of validity generalization. *J Appl Psychol* 1977;62:529-40.
  120. FirthCozens J, Greenhalgh J. Doctors' perceptions of the links between stress and lowered clinical care. *Soc Sci Med* 1997;44:1017-22.
  121. Geer R, Jobs D, Gilfor J, Traber K, Dinges D. Reduced psychomotor vigilance in anesthesia residents after 24-hour call. *Anesthesiology* 1995;83:A1008.
  122. Weinger MB. Vigilance, boredom, and sleepiness. *J Clin Monit Comput* 1999;15:549-52.
  123. Weinger MB, Herndon OW, Gaba DM. The effect of electronic record keeping and transesophageal echocardiography on task distribution, workload, and vigilance during cardiac anesthesia. *Anesthesiology* 1997;87:144-55.
  124. Aronson S, Cook R. Vigilance--a main component of clinical quality. *Anesthesiology* 1998;88:1122-4.
  125. Durnford S. Junior hospital doctors: tired and tested. *BMJ* 1988;297:931-2.
  126. Whang EE, Perez A, Ito H, Mello MM, Ashley SW, Zinner MJ. Work hours reform: perceptions and desires of contemporary surgical residents. *J Am Coll Surg* 2003;197:624-30.
  127. Fischer, JE. Continuity of care: a casualty of the 80-hour work week. *Acad Med* 2004;79:381-3.
  128. Underwood W, Boyd AJ, Fletcher KE, Lypson ML. Executive Committee of the American College of Surgeons-Candidate Associate Society. Viewpoints from generation X: a survey of candidate and associate viewpoints on resident duty-hour regulations. *J Am Coll Surg* 2004;198:989-93.
  129. Egger M, Smith GD. Meta-analysis. Bias in location and selection of studies. *BMJ* 1998;316:61-6.
  130. Akerstedt T, Landstrom U. Work place countermeasures of night shift fatigue. *Int J Industrial Ergonomics* 1998;21:167-78.
  131. Dula DJ, Dula NL, Hamrick C, Wood GC. The effect of working serial night shifts on the cognitive functioning of emergency physicians. *Ann Emerg Med* 2001;38:152-5.
  132. Schwartz RJ, Dubrow TJ, Rosso SE. Guidelines for surgical residents. *Surgery* 1992;127:778-82.
  133. Baldwin DC Jr, Daugherty SR, Tsai R, Scotti MJ. A national survey of residents' self-reported work hours: thinking beyond specialty. *J Acad Med* 2003;78:1154-63.
  134. ACGME, Resident Survey, Data from February/March 2004, reported in "The ACGME's Approach to Limit Resident Duty Hours 12 Months After Implementation: A Summary of Achievements, page 2 and Exhibit 1." Available at: [http://www.acgme.org/acWebsite/dutyHours/dh\\_dutyhoursummary2003-04.pdf](http://www.acgme.org/acWebsite/dutyHours/dh_dutyhoursummary2003-04.pdf). Accessed August 3, 2005
  135. Wu W, Folkman S, McPhee SJ, Lo B. Do house officers learn from their mistakes? *JAMA* 1991;265:2089-94.
  136. Fletcher KE, Davis SQ, Underwood W, Mangrulkar RS, McMahon LF Jr, Saint S. Systematic review: effects of resident work hours on patient safety. *Ann Intern Med* 2004;141:851-7.
  137. Howard DL, Silber JH, Jobs DR. Do regulations limiting residents' work hours affect patient mortality? *J Gen Intern Med*. 2004;19:1-7.
  138. Reason JT. *Human Error*. New York: Cambridge University Press; 1990.
  139. Arnstein F. Catalogue of human error. *Br J Anaesth* 1997;79:645-56.
  140. Ihse I, Haglund U. The Swedish 40-hour work week: how does it affect surgical care. *Surgery* 2003;134:17-8.
  141. Petersen LA, Brennan TA, O'Neil AC, Cook EF, Lee TH. Does housestaff discontinuity of care increase the risk for preventable adverse events? *Ann Intern Med* 1994;121:866-72.
  142. Buysse DJ, Barzansky B, Dinges D, et al. Sleep, fatigue, and medical training: setting an agenda for optimal learning and patient care. *Sleep* 2003;26:218-25.

**APPENDIX 1—List of Studies Included in the Meta-Analysis**

| <b>Authors, Year</b>                      | <b>Type</b> | <b>Subject</b> | <b>N</b> | <b>Sleep Loss</b>         | <b>Performance</b>                                  |
|---|-------------|----------------|----------|---------------------------|---|
| Anderson et al, 2000 <sup>62</sup>        | IG          | Other          | 51       | Chronic Partial           | Vigilance   |
| Angus and Heslegrave, 1985 <sup>63</sup>  | RM          | Other          | 12       | >54 hrs                   | Cognitive function, memory, vigilance               |
| Beaumont et al, 2001 <sup>64</sup>        | RM          | Other          | 16       | >54 hrs                   | Vigilance   |
| Belenky et al, 2003                       | RM          | Other          | 66       | Chronic Partial           | Vigilance   |
| Blagrove et al, 1995 <sup>48</sup>        | IG          | Other          | 14       | Chronic Partial           | Cognitive function, memory, vigilance               |
| Bonnet, 1991 <sup>66</sup>                | IG          | Other          | 44       | <30 hrs                   | Cognitive function, memory                          |
| Bonnet et al, 1995                        | RM          | Other          | 140      | 30–<54 hrs                | Cognitive Function, memory                          |
| Browne et al, 1994 <sup>27</sup>          | RM          | Physician      | 35       | <30 hrs                   | Memory  |
| Caldwell and DeLuc, 1998 <sup>68</sup>    | RM          | Other          | 12       | 30–<54 hrs                | Cognitive function                                  |
| Caldwell and Ramsrott, 1998 <sup>58</sup> | RM          | Other          | 17       | <30 hrs                   | Cognitive function, vigilance                       |
| Casagrande et al, 1997 <sup>69</sup>      | RM          | Other          | 20       | 30–<54 hrs                | Vigilance   |
| Deaconson et al, 1988 <sup>70</sup>       | RM          | Physician      | 26       | <30 hrs                   | Cognitive function, memory, vigilance               |
| DeGennaro et al, 2001 <sup>71</sup>       | RM          | Other          | 8        | 30–<54 hrs                | Vigilance   |
| Dinges et al, 1997 <sup>51</sup>          | RM          | Other          | 16       | Chronic Partial           | Memory, vigilance                                   |
| Doran et al, 2001 <sup>72</sup>           | IG          | Other          | 28       | 30–<54 hrs                | Vigilance   |
| Dorrian et al, 2000 <sup>73</sup>         | RM          | Other          | 18       | < 30 hrs                  | Cognitive function, vigilance                       |
| Drummond et al, 1999 <sup>74</sup>        | RM          | Other          | 13       | < 30 hrs                  | Cognitive function                                  |
| Drummond et al, 2001 <sup>75</sup>        | RM          | Other          | 13       | < 30 hrs                  | Memory  |
| Eastridge et al, 2003 <sup>76</sup>       | RM          | Physician      | 35       | < 30 hrs                  | Clinical performance                                |
| Engel et al, 1987 <sup>77</sup>           | RM          | Physician      | 7        | < 30 hrs                  | Clinical performance                                |
| Fluck et al, 1998 <sup>78</sup>           | RM          | Physician      | 6        | < 30 hrs                  | Memory  |
| Forest and Godbout, 2000 <sup>79</sup>    | IG          | Other          | 18       | < 30 hrs                  | Memory, vigilance                                   |
| Friedman et al, 1971 <sup>1</sup>         | RM          | Physician      | 14       | < 30 hrs                  | Clinical performance                                |
| Gillberg et al, 1994 <sup>80</sup>        | IG          | Other          | 18       | < 30 hrs                  | Vigilance   |
| Gottlieb et al, 1993 <sup>81</sup>        | IG          | Physician      | 86       | 30–<54 hrs                | Cognitive function, memory, vigilance               |
| Grantcharov et al, 2001 <sup>29</sup>     | RM          | Physician      | 14       | < 30 hrs                  | Clinical performance                                |
| Harrison and Horne, 1999 <sup>82</sup>    | IG          | Other          | 20       | 30–<54 hrs                | Memory, vigilance                                   |
| Harrison and Horne, 2000 <sup>83</sup>    | RM          | Other          | 30       | < 30 hrs                  | Cognitive function                                  |
| Hart et al, 1987 <sup>84</sup>            | IG          | Physician      | 30       | < 30 hrs                  | Vigilance   |
| Hawkins et al, 1985 <sup>85</sup>         | IG          | Physician      | 68       | < 30 hrs, Chronic Partial | Cognitive function                                  |
| Heuer et al, 1998 <sup>86</sup>           | RM          | Other          | 8        | >54 hrs                   | Memory  |
| Hockey et al, 1998 <sup>87</sup>          | RM          | Other          | 16       | <30 hrs                   | Cognitive function                                  |
| Horne, 1988 <sup>88</sup>                 | IG          | Other          | 24       | 30–<54 hrs                | Cognitive function                                  |
| Howard et al, 2003 <sup>89</sup>          | RM          | Physician      | 12       | <30 hrs                   | Memory, vigilance, clin. Performance                |
| Jacques et al, 1990 <sup>90</sup>         | RM          | Physician      | 26       | <30 hrs                   | Cognitive function                                  |
| Jaskowski et al, 1997 <sup>91</sup>       | RM          | Other          | 13       | 30–<54 hrs                | Vigilance   |
| Kannan et al, 1997 <sup>92</sup>          | RM          | Physician      | 40       | <30 hrs                   | Cognitive function, memory, vigilance               |
| Kim et al, 2001 <sup>93</sup>             | RM          | Other          | 18       | <30 hrs                   | Cognitive function, memory                          |
| Lamond and Dawson, 1999 <sup>3</sup>      | RM          | Other          | 22       | < 30 hrs                  | Cogn, function, vigilance                           |
| Linde et al, 1999 <sup>94</sup>           | RM          | Other          | 12       | 30–<54 hrs                | Cogn, function, vigilance                           |
| Lingenfelter et al, 1994 <sup>26</sup>    | RM          | Physician      | 40       | < 30 hours                | Cognitive funct., memory, vigilance, clin. Perform. |
| McCarthy and Waters, 1997 <sup>95</sup>   | RM          | Other          | 10       | < 30 hours                | Vigilance   |
| Mak and Spurgeon, 2003 <sup>96</sup>      | RM          | Physician      | 21       | < 30 hours                | Cognitive function                                  |
| Manly et al, 2002 <sup>97</sup>           | RM          | Other          | 8        | < 30 hours                | Memory, vigilance                                   |
| Matsomoto et al, 2002 <sup>98</sup>       | IG          | Other          | 71       | 30 – 54 hours             | Cognitive function, memory, vigilance               |
| Neri et al, 1995 <sup>99</sup>            | RM          | Other          | 10       | < 30 hours                | Vigilance   |
| Pigeau et al, 1995 <sup>100</sup>         | RM          | Other          | 41       | > 54 hours                | Cognitive function                                  |
| Pilcher and Walters, 1997 <sup>101</sup>  | IG          | Other          | 44       | < 30 hours                | Cognitive function                                  |
| Polzella, 1975 <sup>102</sup>             | RM          | Other          | 5        | < 30 hours                | Memory  |
| Reznick and Folse, 1987 <sup>103</sup>    | IG          | Physician      | 33       | < 30 hours                | Cognitive function, memory                          |
| Richardson et al, 1996 <sup>104</sup>     | IG          | Physician      | 13       | < 30 hours                | Memory  |
| Robbins and Gottlieb, 1990 <sup>105</sup> | RM          | Physician      | 23       | 30 – 54 hours             | Cognitive function, memory, vigilance               |
| Steyvers, 1987 <sup>106</sup>             | RM          | Other          | 16       | < 30 hours                | Vigilance   |
| Stone et al, 2000 <sup>107</sup>          | IG          | Physician      | 424      | < 30 hours                | Cognitive function, clin. Performance               |
| Taffinder et al, 1998 <sup>30</sup>       | RM          | Physician      | 6        | < 30 hours                | Clinical performance                                |
| Thomas et al, 2000 <sup>108</sup>         | RM          | Other          | 17       | < 30 hours                | Memory  |
| Webb, 1985 <sup>109</sup>                 | RM          | Other          | 6        | < 30 hours                | Cognitive function, memory, vigilance               |
| Wesensten et al, 2002 <sup>110</sup>      | RM          | Other          | 50       | Chronic Partial           | Memory, vigilance                                   |
| Wimmer et al, 1992 <sup>111</sup>         | IG          | Other          | 25       | < 30 hours                | Cognitive function                                  |
| Wright et al, 1997 <sup>112</sup>         | RM          | Other          | 36       | 30 – 54 hours             | Cogn., function, vigilance                          |

IG refers to independent groups; RM, repeated measures