

ORIGINAL ARTICLE

Sleep deprivation and junior doctors' performance and confidence

K E Lewis, M Blagrove, P Ebden

Postgrad Med J 2002;78:85-87

See end of article for authors' affiliations

Correspondence to:
Dr Philip Ebden; Prince
Philip Hospital, Llanelli,
Dyfed SA14 8QF, UK;
ebdens@aol.comSubmitted 9 April 2001
Accepted 21 August 2002

Purpose of study: To determine whether sleep deprivation affects not only junior doctors' performance in answering medical questions but whether their ability to judge their own performance is also affected by lack of sleep.

Methods: A questionnaire based follow up study in two district general hospitals of the Carmarthenshire NHS Trust. Eleven house officers and 15 senior house officers (SHOs) within the medical directorate participating in the on-call rota were recruited between July 1999 and May 2000.

Results: SHOs answered significantly more questions correctly ($p=0.04$) and were more confident than house officers when they were either correct or incorrect ($p<0.001$). Length of unbroken or continuous sleep is associated with more correct answers ($p=0.03$) and higher energy ($p=0.09$) and confidence ($p=0.07$) scores self rated by the profile of mood states. Length of continuous sleep was not related to the appropriateness of confidence, as measured by the "within-subject confidence-accuracy correlation" ($p=0.919$).

Conclusions: SHOs performed better than house officers even allowing for sleep loss. Sleep deprivation had adverse effects on mood and performance but junior doctors can still monitor their performance and retain insight into their own ability when sleep deprived.

Studies of the effects of sleep deprivation on formal psychological performance in on-call medical staff show conflicting results.¹⁻⁵ Differences in the subjects' specialities, definitions of sleep deprivation, different psychological tests, timing of tests, small numbers of subjects, and different shift patterns may explain this variability.^{6,7} However studies in sleep deprived doctors that examine mood and wellbeing are much more consistent; they usually show adverse effects on mood and well being.^{1,3,7-10}

There are very few studies relating to sleep deprivation and assessment of function of doctors in more relevant tasks to daily work. Some have found deficits in arrhythmia recognition,^{1,8} laparoscopic skills,¹¹ and variable effects on gaining vascular access.¹² We found no studies in the literature about the confidence of doctors in their medical reasoning when they are sleep deprived. This is important because if a doctor were to recognise that their own decision in relation to a medical problem was likely to be incorrect, it would reduce the rate of errors. However, instead, it may be that such insight is itself affected by lack of sleep.

Blagrove and Akehurst found that sleep deprived individuals have low confidence in their incorrect answers to cognitive reasoning tasks, indicating that they appreciate the deficiencies in their performance,¹³ and Dorrian *et al* found that 18 volunteers could still self monitor and predict their performances, even when sleep deprived for up to 28 hours.¹⁴

We have developed a model to look at the effect on junior medical doctors' ability to answer a range of medical questions at various levels of sleep deprivation after a night on-call. We also looked at their trust in their answers when sleep deprived to see whether it was appropriately placed. The outcome measure we use to assess the appropriateness of confidence compares confidence in one's response when that response is correct to confidence in one's response when the response is in fact wrong. The comparison is calculated as the point biserial correlation between accuracy and confidence, and is called the confidence-accuracy correlation. This is described in more detail below. We then calculated whether this correlation was related to their reported maximum length of continuous sleep.

SUBJECTS AND METHODS

Thirty six volunteers (age 23-34 years) were recruited over a 10 month period from the junior medical staff in two district general hospitals in South Wales. They had at least one month's medical experience since qualification. There were no refusals, eliminating one source of (selection) bias. Approval was obtained from the postgraduate dean and local postgraduate organiser. There was a ratio of approximately 2:1 males to females but to ensure confidentiality and the largest number of participants, information on names and sex was not recorded; this was justified because others have found no sex difference in confidence on question items.¹⁵⁻¹⁷

After a night on-call the doctors completed sleep questionnaires, the profile of mood states (POMS), and medical questions between 10 am and 4 pm after being on-call. They were asked to record their bedtime and wake-up time the night before their on-call, to estimate the total number of hours they had slept during their on-call and to estimate their maximum length of continuous sleep during the on-call. Their mood was then self rated by the POMS¹⁸ on two dimensions, energetic-tired, and confident-unsure (this latter mood variable is termed general confidence here so as to distinguish it from confidence in one's response to a question). Each doctor then answered 10 questions (46 stems) selected from medical students' final examination papers and published questions in the style of both parts of the membership examinations for the Royal College of Physicians. Questions consisted of medical multiple choice, electrocardiograms, and data interpretation and were chosen with a range of difficulty to ensure that no individuals obtained 100% correct or 100% incorrect answers. They were not told of the accuracy of their answers during the study. The test lasted 45-90 minutes and there were no time constraints that may affect confidence.

Abbreviations: POMS, profile of mood states; SHO, senior house officer

Table 1 Means of mood, performance, and response confidence for house officers and SHOs

	No of (range) hours total sleep	No of (SD) hours continuous sleep	POMS (SD) general confidence score	POMS (SD) energy score	Mean (SD) confidence when correct	Mean (SD) confidence when incorrect	Mean (SD) No of correct answers	Mean (SD) No of incorrect answers	Mean (SD) No of "don't knows"
House officer	4.8 (0.75–8.0)	3.9 (2.5)	13.1 (3.7)	10.5 (5.6)	3.67* (0.70)	3.16* (0.5)	26.7* (5.0)	11.3 (2.1)	8.0 (5.0)
SHO	3.4 (0.0–6.5)	2.6 (1.4)	14.2 (7.2)	10.9 (8.2)	4.46* (0.4)	3.94* (0.4)	31.6* (5.9)	10.3 (4.3)	4.13 (4.6)

*Two tailed significance at 0.05 level between SHO and house officer.

Range rather than SD accompanies total sleep length because this variable is an approximation by the doctors.

After answering each question the doctors were asked, "How certain are you of your answer?" and rated their confidence in each answer from 1 ("extremely uncertain") to 5 ("extremely certain"). We wanted a measure of whether, if someone is sure of their answer, they are more likely to be correct than if they are less sure of their answer. A standard method of comparing someone's confidence when right to their confidence when wrong is the confidence-accuracy correlation. This is the point biserial Pearson's correlation of the confidence in response to questions (five levels, that is, 1 = "extremely uncertain" to 5 = "extremely certain") with accuracy of response (two levels, that is, correct response and incorrect response). The correlation is calculated for each participant, and can vary from -1.00 to $+1.00$, correlations greater than 0 indicate that confidence in a correct response is higher than confidence in a response that is actually incorrect. In most studies these within-subject correlations are rarely large, but are usually positive. Someone with insight into their performance should be more confident in their answers when those answers are actually correct than when they provide inaccurate answers, such insight is reflected in a higher, more positive correlation, for example $> +0.6$ rather than, for example -0.6 .

The method of statistical analysis then used to determine whether appropriateness of confidence is affected by sleep loss is as follows. A between-subjects correlation was performed between the confidence-accuracy correlation and the maximum length of continuous sleep. This method assesses whether a smaller length of continuous sleep is associated with lower confidence-accuracy correlations. If such an association were found to be significant the doctors would be showing a lack of insight into their poor performance when sleep deprived.

Alternatively, if there was no relationship between each subject's confidence-accuracy correlation and their maximum length of continuous sleep it would indicate that sleep loss does not lead to the lack of insight into accuracy.

(Note that as is usual when a series of correlations is used as data in a further correlational analysis, each individual Pearson r is first converted to Fisher's r' statistic, and the correlation is then run between r' and maximum length of continuous sleep.¹⁹)

Data were analysed using SPSS version 10. Tests were applied for normality and Pearson's correlations were computed. As sleep deprivation is a relative definition we did not arbitrarily divide our participants into sleep deprived versus non-sleep deprived but applied continuous correlations.

RESULTS

Full data were obtained on 26 doctors. Six did not correctly complete their questionnaires and four were rejected because they had not had their normal night's sleep on the night before the on-call. Our group consisted of 15 senior house officers (SHOs) with a median of 61 months (range 13–178 months) since qualification and 11 other house officers with a median of 6.1 months (range 1–15 months) since qualification.

Effects of grade

Table 1 shows the scores for SHOs and house officers on mood, performance, and response confidence. SHOs and house officers' moods post-call were both much lower than the standard population based scores for POMS of 21.9 for general confidence and of 20.11 for energy.¹⁵

SHOs had significantly more correct answers than house officers ($t(24)=2.21$, $p=0.037$) and marginally fewer "don't knows" ($t(24)=2.04$, $p=0.053$). This is reassuring and suggests that the questions were appropriate and reflected the different levels of medical experience. They had a similar number of incorrect responses.

SHOs were significantly more confident than house officers when answering questions correctly ($t(24)=3.57$, $p=0.002$) and when answering incorrectly ($t(24)=4.28$, $p<0.001$), and, with grade as a covariate, doctors were more confident in their answers when correct than when they were incorrect ($F(1,24)=27.54$, $p<0.001$).

Effects of sleep loss when controlling for grade

Table 2 shows a clear trend that maximum length of continuous sleep is positively associated with higher POMS energy ($p=0.09$) and general confidence ($p=0.07$) scores and significantly more correct answers ($p=0.03$). Maximum length of continuous sleep was negatively related to number of "don't knows" but was not associated with the number of incorrect answers. Importantly, however, there was only a very low correlation between maximum length of continuous sleep and the relationship between confidence and accuracy (the confidence-accuracy correlation). This shows that the appropriateness of confidence, that is, its usefulness in distinguishing between correct and incorrect answers, was not affected by sleep loss.

Maximum length of unbroken or continuous sleep correlated more closely with mood scores and correct responses than did length of total sleep. This agrees with anecdotal reporting by doctors that the maximum length of unbroken or

Table 2 Pearson's correlations of length of continuous sleep with mood, performance, and confidence variables (controlling for SHO/house officer grade)

	POMS energy	POMS general confidence	No correct	No incorrect	No don't knows	Confidence when correct	Confidence when incorrect	Within-subject confidence accuracy correlation
Continuous sleep	0.34	0.37	0.44	-0.18	-0.39	-0.31	-0.34	-0.02
p Value (df=23)	0.09*	0.07*	0.03**	0.4	0.06*	0.127	0.097*	0.919

*Two tailed significance at 0.1 level.

**Two tailed significance at 0.05 level.

continuous sleep is more important to their sense of wellbeing than total sleep length.

DISCUSSION

This study suggests that junior doctors undergoing sleep deprivation when on-call, have significant sleepiness and lowered general confidence. They answer significantly fewer medical questions correctly, but they do not have inappropriate confidence in their answers, in that they have higher confidence in their correct than in their incorrect answers, and the relationship between confidence when right and confidence when wrong (the confidence-accuracy correlation) is not affected by amount of sleep loss. In fact, doctors undergoing sleep deprivation are more likely to answer "don't know" to medical answers. We thus conclude that, whereas sleep deprived doctors are more likely to make mistakes, they do still have insight into their deficient performance. This may ameliorate possible adverse consequences, in that these doctors when sleepy may then deliberately take longer over a task, or recognise that they need to ask for a second opinion.

A strength of the study is its use of medical knowledge, whereas many sleep deprivation studies use simpler speed of cognition, reaction time, or vigilance measures. We also excluded subjects who did not have their normal night's sleep before their "on-call" started. Any prior abnormal night's sleep could have carry over effects. A weakness is the lack of a baseline version of the questions for each doctor before any loss of sleep, which would have enabled a within-subjects analysis to be performed. It is difficult to choose questions of equal difficulty and we did not want to repeat the same questions in a normal and sleep deprived state to avoid any learning effect. We also accept that effects of sleep loss on performance may have been different if distracter questions had been used, or participants had been under time pressure to complete the answers. Although distracters and time pressures do exist in real life, we wanted to ensure that all subjects had as many data points as possible for calculating the within-subjects correlation. We therefore needed participants to attempt all the questions and not leave any out due to time limitations. This also met our aim of participants having similar numbers of data points in their within-subjects confidence-accuracy correlations. Furthermore, although we had some interest in performance from our doctors, they were not told individual scores to avoid competitiveness and as little extra stress as possible. Our main aim was the study of confidence in performance, and the self assessment of this confidence may also have been aided by the lack of time pressure.

In this cohort of junior doctors, sleep deprivation leads to fewer correct answers to medical questions and lowered mood states, but they had appropriately low confidence when wrong, even after sleep loss, and this suggests they could still appreciate their deficiencies when sleep deprived.

ACKNOWLEDGMENT

We are grateful to the junior doctors who took part in our survey; Diana Morgan, for helping to prepare the manuscript; Ann Leeuwerke, for helping with library literature searches; and Siân Morris and Sue Harrison in the Postgraduate Centre for helping to recruit and supervise the doctors' completion of questionnaires.

.....

Authors' affiliations

K E Lewis, Morriston Hospital, Swansea, Wales
M Blagrove, University of Wales Swansea, Swansea, Wales
P Ebdon, Prince Philip Hospital, Llanelli, Wales

REFERENCES

- 1 **Lingenfeller TH**, Kaschels R, Weber A, *et al*. Young hospital doctors after night duty: their task specific cognitive status and emotional condition. *Med Educ* 1994;**28**:566-72.
- 2 **Rubin R**, Orris P, Lau SL, *et al*. Neurobehavioural effects of the on-call experience in house staff physicians. *J Occup Med* 1991;**33**:8-13.
- 3 **Orton DI**, Gruzeliier JH. Adverse changes in mood and cognitive performance of house officers after night duty. *BMJ* 1989;**298**:21-3.
- 4 **Binks PG**, Water WF, Hurry M. Short-term total sleep deprivation does not selectively impair cortical functioning. *Sleep* 1999;**22**:328-34.
- 5 **Richardson G**, Wyatt KJ, 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.
- 6 **Kapur N**, House A. Job satisfaction and psychological morbidity in medical house officers. *J R Coll Physicians Lond* 1997;**31**:162-7.
- 7 **Leung L**, Becker CE. Sleep deprivation and house staff performance. Update 1984-1991. *J Occup Med* 1992;**34**:1153-60.
- 8 **Friedman RC**, Bigger JT, Kornfield DS. The intern and sleep loss. *N Engl J Med* 1971;**285**:201-3.
- 9 **Bartle EJ**, Sun JH, Thompson L, *et al*. The effects of acute sleep deprivation during residency training. *Surgery* 1988;**104**:311-16.
- 10 **Engel W**, Seine R, Powell V, *et al*. Clinical performance of interns after being on-call. *South Med J* 1987;**80**:761-3.
- 11 **Reznick RK**, Folse JR. Effect of sleep deprivation on the performance of surgical residents. *Am J Surg* 1987;**154**:520-5.
- 12 **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.
- 13 **Blagrove M**, Akehurst L. Effects of sleep loss on confidence-accuracy relationships for reasoning and eyewitness memory. *J Exp Psychol Applied* 2000;**6**:59-73.
- 14 **Dorrian J**, Lamond N, Dawson D. The ability to self-monitor performance when fatigued. *J Sleep Res* 2000;**9**:137-44.
- 15 **Ross JA**, Hogaboam AG, Hannay L. Predictors of teachers' confidence in their ability to implement computer-based instruction. *J Educ Comput Res* 1999;**21**:75-97.
- 16 **Lundeberg MA**, Fox PW, Brown AC, *et al*. Cultural influences on confidence. *J Educ Psychol* 2000;**92**:152-9.
- 17 **Sleeper LA**, Nigro GN. Not who you are but who you're with: self-confidence in achievement settings. *Sex-Roles* 1987;**16** [1-2]:57-69.
- 18 **Lorr M**, McNair DM. *Profile of mood states [bi-polar form—POMS-BI]*. San Diego, California: Educational and Industrial Testing Service, 1984.
- 19 **Howell DC**. *Statistical methods for psychology*. 2nd Ed. Boston: PWS-Kent, 1987.