

29 Gough DA, Lucisano JV, Tse PHS. Two-dimensional enzyme electrode sensor for glucose. *Anal Chem* 1985;57:2351-7.

30 Pickup JC, Shaw GW, Claremont DJ. Potentially-implantable, amperometric glucose sensors with mediated electron transfer: improving the operating stability. *Biosensors* (in press).

31 Claremont DJ, Sambrook IE, Penton C, Pickup JC. Subcutaneous implantation of a ferrocene-mediated glucose sensor in pigs. *Diabetologia* 1986;29:817-21.

32 Gordon GS, Moses AC, Silver RD, Flier JS, Carey MC. Nasal absorption of insulin: enhancement by hydrophobic bile salts. *Proc Natl Acad Sci USA* 1985;82:7419-23.

33 Salzman R, Manson JE, Griffing GT, *et al*. Intranasal aerosolized insulin. Mixed meal studies and long-term use in type 1 diabetes. *N Engl J Med* 1985;312:1078-84.

34 McMartin C, Hutchinson LEF, Hyde R, Peters GE. Analysis of structural requirements for the absorption of drugs and macromolecules from the nasal cavity. *J Pharm Sci* 1987;76:535-40.

35 Frauman AG, Cooper ME, Parsons BJ, Jerums G, Louis WJ. Long-term use of intranasal insulin in insulin-dependent diabetic patients. *Diabetes Care* 1987;10:573-8.

36 Lassman-Vague V, Thiers D, Vialettes B, Vague P. Preprandial intranasal insulin. *Lancet* 1988;i:367-8.

Junior hospital doctors: tired and tested

Should the rules for pilots apply to doctors?

Few people will be surprised that a paper to be published soon in the *BMJ* will show cognitive impairment and mood changes in junior doctors working long hours. We would expect such findings in subjects deprived of sleep from other walks of life, and we might condemn their work schedules, but doctors have traditionally felt free to ignore the advice they give to others. Recently, however, interest in whether tired doctors might be dangerous has reached a level that is difficult to ignore, and the BMA's annual meeting has called for limits on the hours that junior staff work (16 July, p 217).

Those doctors who consider that the houseman's year is a necessary evil, a rite, or even an advantageous education^{1,2} might reasonably ask whether reduced performance on cognitive tests automatically implies reduced medical ability. Research on this subject, although sparse, goes back a long way, and the chance findings of Poulton *et al* in 1978 are a good example of how mixed the results may be.³ In their study on doctors deprived of sleep they found that performance of a grammatical reasoning task was significantly different if the test was immediately followed by a second task of sorting laboratory results into normal and abnormal. Those doctors who did only the grammatical reasoning task showed a reduced performance once they had accumulated three hours of sleep deficit (over the previous one to three days), whereas those who did both tests showed no decline in their performance until their sleep deficit reached eight hours.³

That the research findings may be confusing is not surprising. We all know that we perform poorly if we are almost asleep, but neither are we at our best when overexcited or panicking. Somewhere in between is the "optimum level of arousal," originally described in the inverted U curves of Yerkes and Dodson in 1908.⁴ Factors that shift our arousal up or down abound in the environment as well as within ourselves. The effect of noise and heat, for example, are well described⁵ as are those of natural circadian rhythms⁶ and coffee.⁷ There are also influences from the tasks performed: the duration of the task,⁸ feedback on performance,⁸ standards of performance learnt at the time the task itself was learnt,³ and the complexity of the task⁹ are all important. Even when factors such as these have been carefully balanced out in the experimental design there is still the uncomfortable feeling that arousal may be partly under volitional control. Several studies have shown that motivation has a strong influence on performance.⁹⁻¹¹

Despite this multitude of variables there is a considerable consensus on the effects of long work hours and reduced sleep—at least in the laboratory. In general, the earliest changes are fatigue and alterations of mood. These effects should not be ignored as they may influence interpersonal relationships and the efficiency of a team. Detectable changes in individual performance appear later, and Nicholson and Stone summarise these in saying "the effects of sleep loss

vary widely. . . Interesting tasks with relatively simple motor skills are resistant for periods as long as 60 hours, but routine monotonous tasks show a rapid and severe decrement after 18 hours without sleep. . . . Sleep deprivation leads to brief intermittent lapses in performance which increase in frequency and duration, and so impaired performance involves missed signals and errors of omission."¹²

So how do people doing real jobs perform when tired? Evidence from aviation and military workers shows that real tasks may be less affected than simulated tasks—for example, one night's loss of sleep had a greater effect on flying a simulator than on flying a real aircraft, although both were significantly affected.¹³ Many of the real world studies might be less sensitive than laboratory studies because of the problems of experimental control, but different degrees of motivation are probably also important.

If performance at a genuine job tends to be preserved then this might explain why there have been so few reports of tired doctors giving inadequate care. An alternative explanation is that such cases may be underreported. The Civil Aviation Authority fund the Royal Air Force Institute of Aviation Medicine to run a confidential human factors incident reporting programme, to which any pilot or air traffic controller may write without fear of identification or disciplinary action. About a third of the incidents reported result from sleep, fatigue, or roosting, and many could easily have been accidents.¹⁴ Although there have been few aircraft accidents actually attributed to fatigue, it would be foolish to assume that fatigue does not affect pilots. Is it not then foolish to assume that fatigue does not affect junior doctors?

The differences in the rules for pilots and doctors are interesting (p 938). In most cases the duty periods of pilots are limited to eight to 14 hours. Aircraft captains may use their discretion to increase this on special occasions, and extensions may also apply if crew members are able to take rest in flight. If a bunk is available, for example, the flying duty period may be increased by half the time spent in the bunk—to a maximum of 18 hours total duty. After any flying duty aircrew must rest for at least 12 hours, and the rest must not be shorter than the preceding duty period.¹⁵ In comparison, about two thirds of junior doctors work rotas of one in two or one in three—and 5% of those on call all night work without a break from midnight to 8 am as well as working the day shifts each side.¹⁶

Both flying and medicine are unforgiving of seemingly minor slips. Although fatigued staff may perform surprisingly well at a motivating challenge, they are likely to make an increasing number of minor errors in subsidiary tasks. A wrong decimal point, a forgotten drug interaction, or incorrect labelling of right or left and the results may be catastrophic. To reduce the hours worked by junior doctors, on the other hand, may be difficult and unpopular as it would imply an

increase in either the number of medical graduates or the length of their postgraduate training.

These are controversial matters with financial and professional implications. Junior doctors have been campaigning against their long working hours, and the BMA has been negotiating with the health departments to reduce these hours (p 937). But unless the NHS comes up with a solution soon consumer opinion will force one on them—as has already happened in the United States (p 938).¹⁷ After all, would you like to put your well being in the hands of a pilot who has been working without rest for 30 hours or more?

SIMON DURNFORD

Consultant in Aviation Medicine (Army),
Royal Air Force Institute of Aviation Medicine,
Farnborough,
Hampshire GU14 6SZ

The views expressed in this article are those of the author only and should not be assumed to reflect the opinions of either the Ministry of Defence (Army) or the Royal Air Force Institute of Aviation Medicine.

1 Jacoby M. A father's letter to a new intern. *JAMA* 1981;245:1031.

2 Leitzell JD, Turkewitz LJ, Ratnoff OD, et al. Internship: physicians respond to Norman Cousins. *JAMA* 1981;246:2141-3.

- 3 Poulton E, Hunt G, Carpenter A, Edwards R. The performance of junior hospital doctors following reduced sleep and long hours of work. *Ergonomics* 1978;21:279-95.
- 4 Yerkes K, Dodson J. The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative and Neurological Psychology* 1908;18:459-82.
- 5 Poulton E. *Environment and human efficiency*. Springfield, Illinois: Thomas, 1970.
- 6 Klein K, Wegmann H. *Significance of circadian rhythms in aerospace operations*. Neuilly sur Seine, France: NATO Advisory Group for Aerospace Research and Development, 1980. (AGARDOgraph No 247.)
- 7 Dews P, ed. *Caffeine*. New York: Springer Verlag, 1984:86-103.
- 8 Johnson L, Naitoh P. *The operational consequences of sleep deprivation and sleep deficit*. Neuilly sur Seine, France: NATO Advisory Group for Aerospace Research and Development, 1976. (AGARDOgraph No 193.)
- 9 Angus R, Heselgrave R. The effects of sleep loss and sustained mental work: implications for command and control. In: *Sustained intensive air operations: physiological and performance aspects*. Neuilly sur Seine, France: NATO Advisory Group for Aerospace Research and Development, 1983:11.1-11.21. (AGARD CP No 338.)
- 10 Haslam D. The incentive effect and sleep deprivation. *Sleep* 1983;6:362-8.
- 11 Kleinhans G. Sustained military operations with particular reference to prolonged exercise. In: *Sustained intensive air operations: physiological and performance aspects*. Neuilly sur Seine, France: NATO Advisory Group for Aerospace Research and Development, 1983:30.1-30.9. (AGARD CP No 338.)
- 12 Nicholson A, Stone B. *Sleep and wakefulness handbook for flight medical officers*. 2nd ed. Neuilly sur Seine, France: NATO Advisory Group for Aerospace Research and Development, 1987. (AGARDOgraph No 270[E].)
- 13 Farmer E, Green R. The sleep deprived pilot: performance and EEG response. In: Sorsa M, ed. *Report of the sixteenth conference of WEEAP*. Helsinki: Finnair Training Centre, 1985: 155-62.
- 14 Green R, Skinner R. CHIRP and fatigue. *The Log* 1987;48:6-11.
- 15 Civil Aviation Authority. *The avoidance of excessive fatigue in aircrews*. 2nd ed. London: CAA, 1982. (CAP 371.)
- 16 Office of Manpower and Economics. *Appendix E to seventeenth report of the Review Body on Doctors' and Dentists' Remuneration*. London: HMSO, 1987. (Cmnd 127.)
- 17 Asch D, Parker R. The Libby Zion case. One step forward or two steps backward? *N Engl J Med* 1988;318:771-5.

Controversy over mammography screening

It should save lives

In 1987 amid much political activity the government accepted the recommendations of the Forrest committee and announced that all women between 50 and 64 would be offered mammography every three years in a national screening campaign.¹ Doubts have since surfaced about the programme, and on p 971 Skrabanek outlines his case against national screening. What might be considered to be the establishment view is put forward by Warren on p 969.

Modern mammography detects breast cancer long before it may be palpated by the surgeon. The impetus to screen for breast cancer with mammography came from two large retrospective, randomised, and controlled trials published in the early 1980s. These studies from Sweden² and the United States³ showed that screening produced a 30% reduction in mortality that was significant in women aged 50-65. Others have suggested that the benefit in both trials may be even greater.^{4,5} In the American trial the screened group included a third of women who were offered screening but did not take it up as well as others who dropped out. Women in the American trial were offered mammography with two views and clinical examination yearly. The Swedish trial differed in that only one view was offered every two to three years and there was no clinical examination. This regimen is comparable with that recommended by the Forrest report. Skrabanek questions the conclusions from these trials and argues that the Forrest report is a consensus document that does not mention the arguments of the dissenting minority. If this is true then the dissenters have been notable by their silence elsewhere. One exception quoted by Skrabanek is Wright, a Canadian surgeon who criticised the evidence from the original American trial.⁶ He claimed that there were 6% more deaths from all causes in screened women compared with those in the controls. In the correspondence that followed he admitted, however, that this was not the case and that he had made a miscalculation.^{4,7,8} The question of the "slightly higher" overall mortality in the Swedish trial

remains unanswered; Skrabanek gives no figure, but Wright after communicating with one of the original authors states that it is only 1%.⁸ Most experts think that the evidence from these trials is strong despite there being anomalies when small subgroups are examined.⁹ Feig recently reviewed the data from the five main trials of mammography screening^{2,3,10-12} and concluded that yearly two view mammography with a physical examination in women from 40 onwards could reduce mortality by at least 40% and possibly by as much as 50%.⁵ Skrabanek's statement that the yearly benefit would be one death for every 15 000 women does not match up with the figures from the trials.⁶

The critics of mammography will think that their case is supported by a third prospective and randomised trial published today (p 943). This study from Malmö in Sweden offered women over 45 five rounds of screening at intervals of 18 to 24 months. When the trial ended—after nearly nine years—there had been no overall fall in mortality in the group offered screening. But among women over 55 mortality fell by a fifth in women who were screened despite a lower rate of acceptance among the older women than the younger women. Furthermore, mortality fell in the final years of the trial and just after it ended in both the whole group offered screening and those over 55. Women under 55 did not show any fall.

For every 1000 women screened for the first time about five to seven will be shown to have cancer. Although modern mammography has a sensitivity of about 80% and a specificity of about 95%, of much greater importance is the positive predictive value—true positive results divided by true positive and false positive results.¹³ Warren, Skrabanek, and Wright⁶ are concerned about the positive predictive value, but it is not even mentioned in the Forrest report. Skrabanek cites the Canadian national breast screening study,¹⁴ in which the average positive predictive value from five centres was 8.6%—that is, after mammography it had nearly 11 false positive results for every true positive result. In an American study the