

even if it is true on site, the nature of the dusts involved and the quantities affecting the local population could well be different.

Afacan's allusion to research into occupational exposures in the mining industry is irrelevant in the community setting and his reliance on occupational exposure standards is unwise. The numbers of people employed in open cast mining is low, so it is unlikely to improve community health by bringing jobs and increased prosperity to the area. Chambers has eloquently expressed the feelings that the "spectre of open cast mining" brings to a community.<sup>1</sup>

We would welcome a major epidemiological study to evaluate the size of the risk and the nature of the mechanisms involved, but in 1989 the Secretary of State for Wales felt that no such study was required.

J M F TEMPLE

Glynneath,  
West Glamorgan SA11 5AL

A M SYKES

University College,  
Swansea

1 Chambers J. Paradise postponed. *BMJ* 1988;297:1688-90.

## Exercise, fitness, and health

EDITOR,—Daphne Glog's editorial confirms what all exercise enthusiasts have been saying for years: that exercise is good for us and that most people do not take enough.<sup>1</sup> Glog concludes that all doctors should ask about exercise when they see patients and advise on suitable exercise and local facilities. Doctors may also, however, unwittingly reduce some patients' capacity to take exercise by prescribing drugs that may interfere with exercise metabolism. Endurance exercise has specific benefits by favourably modifying lipid profiles.<sup>2</sup> Any endurance exercise of moderate intensity (about 50% maximal oxygen uptake) relies heavily on oxidation of fat for resynthesis of ATP and production of energy, thereby preserving carbohydrate stores. Drugs that modify the metabolism of fat and reduce the availability of fatty acids for oxidation would therefore be expected to have an impact on a patient's capacity to take exercise. Our recently completed trial suggests that a lipid lowering drug (a derivative of nicotinic acid) considerably reduced oxidation of fat, a fibrate had an intermediate effect, and a 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitor had no effect.

A suitably designed exercise regimen is often prescribed as part of a rehabilitation programme for patients suffering from hypertension or angina or recovering from myocardial infarction. Concurrent drug treatment for such patients may include a  $\beta$  blocker, which has a cardioprotective effect.<sup>1</sup> Our observations suggest that in comparable doses  $\beta_1$  selective agents may have less of an impact on total oxidation of fat than non-selective agents. During one hour's exercise the mean proportion of energy supplied by fat in eight healthy men was 41.5% (SD 10.9%) with a placebo, 34.2% (10%) with a  $\beta_1$  selective drug, and 28.6% (8.2%) with a non-selective drug. Furthermore,  $\beta_1$  selective drugs reduce the unwanted  $\beta_2$  mediated effects on glucose metabolism and peripheral vasoconstriction associated with non-selective drugs and should therefore have less of an impact on exercise metabolism. Doctors might also give advice on the timing of exercise as with most  $\beta$  blockers peak plasma concentrations are reached one to three hours after an oral dose, with exercise being more difficult during this period.

Thus care should be taken that concurrent prescription of drugs and exercise are not in conflict. A patient taking a non-selective  $\beta$  blocker and a fibrate (a not uncommon combination) might have considerable difficulty in taking modest

exercise such as a brisk walk for one to two hours. Further studies are needed to find therapeutic regimens that reduce patients' risk of suffering from the effects of coronary artery disease but also allow exercise which may have the same goals and, in addition, improve the quality of life.

MARTIN J KENDALL  
ANTHONY HEAD

Clinical Investigation Unit,  
Department of Medicine,  
Queen Elizabeth Hospital,  
Birmingham B15 2TH

- 1 Glog D. Exercise, fitness, and health. *BMJ* 1992;305:377-8. (15 August.)
- 2 Dufaux B, Assman G, Hollman W. Plasma lipoproteins and physical activity. *Int J Sports Med* 1982;3:123-36.
- 3 Olsson G, Wikstrand J, Warnold I, Manger Cats V, McBoyle D, Herlitz J, et al. Metoprolol-induced reduction in postinfarction mortality: pooled results from five double blind randomised trials. *Eur Heart J* 1992;13:28-32.

EDITOR,—Daphne Glog believes that there is no dispute that exercise is beneficial.<sup>1</sup> Although over 800 deaths have occurred in accidents during sports and leisure exercise in the last five reported years (1986-90)<sup>2</sup> as well as other deaths during exercise from pre-existing disease, the benefits of exercise with regard to mortality are well established, at least from the public health point of view. From this vantage, discounting the future health gains derived from fitness to set against the present hazards may not be appropriate.<sup>3</sup> Last Saturday, however, when I played cricket in a team without its regular wicket keeper, who was in hospital recovering from a heart attack that occurred during a game of squash, against a team missing a player who had died after being struck by lightning during a cricket match, the balance was not so clear.

Furthermore, taking morbidity into account leaves matters still less clear cut. In 1988 over half a million people aged 16 and over and 200 000 children attended accident and emergency departments in the United Kingdom after sports accidents occurring outside the home (unpublished data from home accident surveillance system 1988, Consumer Safety Unit of Department of Trade and Industry). In 1987 the general household survey included 184 adults who had had sports accidents that resulted in contact with a doctor.<sup>4</sup> These patients took 670 days off work as a result. Translating these small numbers into national estimates yields an estimated 1.5 million injuries related to exercise needing contact with a doctor and 5.5 million days lost from work. A report, commissioned by the Sports Council of England, on a national survey of injuries and illnesses related to sports and exercise in 17 500 adults, which should be published later this year, shows that even these numbers underestimate the problem. In addition, many non-trivial injuries occur which are treated by physiotherapists, dentists, complementary therapists, nurses, and others, and there is the possibility of chronic conditions developing later in life.

Thus as well as indisputably conferring health benefits, exercise indisputably incurs substantial health costs. As always, the trick is to weigh these carefully against each other to determine whether exercise is, in fact, beneficial.

JON NICHOLL

Medical Care Research Unit,  
Department of Public Health Medicine,  
Medical School,  
Sheffield S10 2RX

- 1 Cloag D. Exercise, fitness, and health. *BMJ* 1992;305:377-8. (15 August.)
- 2 Office of Population Censuses and Surveys. Fatal accidents occurring during sporting and leisure activities 1986-1990. London: HMSO, 1988-91. (DH4 88/3, 88/6, 89/4, 90/4, 91/4.)
- 3 Discounting health care: only a matter of timing [editorial]. *Lancet* 1992;340:148-9.
- 4 Office of Population Censuses and Surveys. *General household survey, 1987 data*. London: HMSO, 1990.

## Validating the SF-36

EDITOR,—We would like to draw attention to a number of errors and anomalies of interpretation in Brazier and colleagues' paper on validating the SF-36 health survey questionnaire.<sup>1</sup>

Firstly, it is inappropriate to attempt to validate a health survey questionnaire (the SF-36) against a measure of perceived distress (the Nottingham health profile). In addition to measuring different concepts, the items in the two measures were generated by different methods. In the SF-36, the items represent professionals' assumptions about issues of relevance to health status. The items in the Nottingham health profile were derived directly from patients and other lay people, describing their perceptions of ill health. It has also been made clear in the major publication on the measure that the Nottingham health profile is not suitable for use in general population or general practice surveys. The instrument works best with selected groups, particularly chronically ill and elderly patients.<sup>2</sup> Furthermore, presenting two similar measures together is likely to lead to response bias.

Secondly, the paper does not report one of the Nottingham health profile's sections at all: sleep. Since sleep problems are common in ill health this would appear to be an important omission from the SF-36, possibly resulting from the method by which items were generated.

Thirdly, the authors claim that the test-retest values obtained for the SF-36 indicate excellent reliability. This is in fact incorrect, as the values are lower than those generally accepted as indicating that a scale is reliable.<sup>3</sup>

Lastly, if a questionnaire cannot be shown to be reliable then there is little value in testing its validity. However, even in the authors' own terms the validity of the SF-36 was not shown. The socioeconomic class gradients are not as consistent as would be expected on several of the dimensions. Two of the subscales do not correspond to the presence of chronic physical problems, and the correlations between dimensions do not accord with logic (vitality and mental health correlated 0.69).

A close inspection of the data presented in the paper suggests that they do not support either the reliability or the validity of the SF-36 in the United Kingdom. It is noteworthy that nothing has been published on the SF-36 in its country of origin.

We welcome this attempt to find a much needed measure for assessing health in the general population. However, the choice of a measure, the techniques by which it is tested, and the interpretation of data require more careful attention.

SONJA M HUNT  
STEPHEN P MCKENNA

Care Outcomes Group,  
Rheumatology and Rehabilitation Research Unit,  
School of Medicine,  
University of Leeds,  
Leeds LS2 8NZ

- 1 Brazier JE, Harper R, Jones NMB, O' Cathain A, Thomas KJ, Usherwood T, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160-4. (18 July.)
- 2 Hunt SM, McEwen J, McKenna SP. *Measuring health status*. London: Croom Helm, 1986.
- 3 McDowell I, Newell C. *Measuring health: a guide to rating scales and questionnaires*. New York: Oxford University Press, 1987.

EDITOR,—It is vital that the validity and reliability of health status measures are thoroughly investigated. Test-retest reliability is particularly important because measurement error can make it more difficult to detect real differences between populations. Although Brazier and colleagues have carried out a large and well designed evaluation of their version of the SF-36 health survey questionnaire,<sup>1</sup> we found their analysis of test-retest reliability incomplete. This aspect of their paper is highlighted by the claim of a level of test-retest

repeatability of only one scale point on a range 0-100, which seems unlikely for a score derived from responses to questions concerning subjective symptoms. They say they have used the method of Bland and Altman,<sup>2</sup> but the results given—the mean difference and the distribution of the differences—are not good measures of reliability. Moreover, Bland and Altman proposed that the limits of agreement should be used to assess repeatability. For example, twice the standard deviation of the differences gives a range within which 95% of the differences will lie.

A mean difference of zero implies that there is no consistent trend affecting test and retest results (all the patients getting better, for example). It does not imply that the measure is reliable. The table gives some hypothetical data to show how you can get a mean difference of 0 and still have widely different test and retest scores, because differences in different directions cancel each other out whatever their absolute size.

#### Obtaining a mean difference of zero

Subject	Test	Retest	Difference
A	50	80	+30
B	30	20	-10
C	60	40	-20
D	70	70	0
Mean difference			0

Bland and Altman explained that the distribution of the differences can be assumed to be normal because they reflect measurement error rather than variations between subjects. Confirming this has no bearing on the question of reliability.

Brazier and colleagues have not supplied data on the limits of agreement. Thus their conclusion that test-retest reliability was excellent is unsupported. To permit an estimate of the limits of agreement, we would ask the authors to present the standard deviations of the differences and the repeatability coefficients for the six dimensions of the scale along with comparable repeatability data for the Nottingham health profile.

PATRICK GOMPERTZ  
ROWAN HARWOOD  
SHAH EBRAHIM  
EDWARD DICKINSON

Department of Health Care of the Elderly,  
Royal London Hospital (Mile End)  
London E1 4DG

- Brazier JE, Harper R, Jones NMB, O'Cathain A, Thomas KJ, Usherwood T, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160-4. (18 July.)
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical assessment. *Lancet* 1986;i:307-10.

**AUTHORS' REPLY.**—We are glad to be able to clarify the issues raised by Sonja M Hunt and Stephen M McKenna and Patrick Gompertz and colleagues.

Despite the fact that its developers now claim the Nottingham Health Profile to be inappropriate for use in a general population, it is used in this way.<sup>1,3</sup> Hunt and McKenna have themselves used the profile in community and general practice settings<sup>4</sup> and have in the past advocated its use in epidemiological studies.<sup>4</sup> A recently published review of health status measures has suggested that the Nottingham Health Profile is appropriate for use in general populations.<sup>5</sup> The profile is widely regarded as a well established instrument for the measurement of perceived health status in the United Kingdom and it was the obvious choice of instrument to compare with the SF-36.

Both letters raise important questions concerning test-retest reliability. Our study investigated the performance of the SF-36 in a general population and not at the individual level. For this

purpose the Bland-Altman technique<sup>7</sup> is an essential addition to correlation coefficients for examining the reliability of an instrument. The Bland-Altman technique looks at the distribution of the differences between the test and retest scores, requiring them to be normally distributed with mean zero. Bland and Altman recommend that the plots of the differences between test and retest scores be visually examined to detect bias. However, there was not sufficient space to present an additional eight figures, and thus we restricted our presentation to the mean differences and their significance for the SF-36 dimensions and the proportion of observations that were within the 95% confidence interval. Our results clearly show that the test-retest score differences of the SF-36 are approximately normally distributed with a mean close to zero, and they indicate that the test and retest score distributions are nearly identical and thus, in conjunction with the correlation coefficients, suggest that the SF-36 reliably measures population scores. Gompertz and colleagues have asked us to present the standard deviations of the differences, and these are presented in the table along with the 95% confidence intervals of the distributions of the differences.

#### Means, standard deviations, and 95% confidence intervals of distributions of difference in test-retest scores by dimension

Dimension	Mean	Standard deviation	95% Confidence interval
Physical functioning	0.49	4.92	-9.35 to 10.33
Social functioning	0.15	1.47	-2.79 to 3.09
Role limitations (physical)	0.57	3.86	-7.15 to 8.29
Role limitations (emotional)	0.44	2.91	-5.38 to 6.26
Pain	0.39	1.37	-2.35 to 3.13
Mental health	0.70	3.05	-5.40 to 6.80
Vitality	0.71	2.64	-4.57 to 5.99
General health perception	0.40	3.54	-6.68 to 7.48

The socioeconomic gradients we observed are exactly as would be expected, with a consistent contrast between the SF-36 scores for the higher two classes and those for the bottom two groups. The variability observed in the two intermediate classes is similar to that shown across a range of conditions in general practice.<sup>8</sup> Similarly the high correlation between mental health and vitality scores is entirely consistent with clinical expectations; poor mental health is associated with low vitality. There was no inconsistency between the mean scores in any dimension and the presence or absence of chronic physical problems.

Two papers by the American developers of the SF-36 have recently been published.<sup>9,10</sup>

We thank Hunt and McKenna for their contribution to the debate regarding the usefulness of the SF-36 but wish to restate that our paper clearly shows acceptable reliability and validity of the SF-36 for primary care and community surveys. It remains necessary to assess its usefulness in a range of clinical settings.

JOHN BRAZIER  
ROSEMARY HARPER  
NICOLA JONES  
KATE THOMAS  
LINDA WESTLAKE

Medical Care Research Unit,  
Department of Public Health Medicine,  
University of Sheffield Medical School,  
Sheffield S10 2RX

TIM USHERWOOD

Department of General Practice,  
University of Sheffield Medical School,  
Sheffield S10 2RX

- Bucquet D, Curtis S. Socio-demographic variation in perceived illness and the use of primary care: the value of community survey data for primary care service planning. *Soc Sci Med* 1986;23:737-44.
- Wilkin D, Leavey R. A comparison of two survey measures of health status. *Soc Sci Med* 1988;27:269-75.

- Baum FE, Cooke RD. Community health needs assessment: use of the NHP in an Australian study. *Med J Aust* 1989;150:581-90.
- Hunt SM, McEwen J, McKenna SP. Measuring health status: a new tool for clinicians and epidemiologists. *J R Coll Gen Pract* 1985;35:185-8.
- Hunt SM, McKenna SP, McEwen J, Williams J, Papp E. The Nottingham health profile: subjective health status and medical consultations. *Soc Sci Med* 1981;15A:221-9.
- Bowling A. *Measuring health: a review of quality of life measurement scales*. Milton Keynes: Open University Press, 1991.
- Bland JM, Altman DG. Statistical methods of assessing agreement between two methods of clinical measurement. *Lancet* 1986;i:307-10.
- Royal College of General Practitioners, Office of Population Censuses and Surveys, Department of Health. *Morbidity statistics from general practice 1981-1982*. London: HMSO, 1990. (Series MBS No 2.)
- Ware EJ, Sherbourne CD. The MOS 36-item short-form health survey (SF-36) conceptual framework and item selection. *Medical Care* 1992;30:473-81.
- McHorney CA, Ware JE, Rogers W. The validity and relative precision of MOS short- and long-form health status scales and Dartmouth COOP charts. *Medical Care* 1992;30(suppl):ms 253-65.

## Respiratory medicine: the casualties

**EDITOR.**—As those most directly affected by the Joint Planning Advisory Committee's recent freeze on senior registrar appointments in respiratory medicine we agree with the views expressed in the letter from a group of consultants in respiratory medicine.<sup>1</sup> A group has now been formed to represent career and research registrars in respiratory medicine.

We have conducted a nationwide survey of clinical and research registrars in respiratory medicine; so far we are aware of 53 registrars. This figure is much greater than previously thought. Of these, 41 have completed or are working towards an MD or a PhD in respiratory medicine and half of these have been in the specialty three or more years. Thus an enormous commitment to the specialty has already been made. Most of these registrars are over 30 years old and in the current system have no realistic chance of retraining in another specialty.

As well as the disastrous effect on the people concerned this freeze has catastrophic implications for the future of the specialty. We believe that applying a simple mathematical formula to a complex problem is inappropriate. We have written to the chairman of the Joint Planning Advisory Committee urging that the freeze be lifted immediately pending a comprehensive discussion of the wider issues.

PAUL DILWORTH

Charing Cross Hospital,  
London W6 8RF

TIM HOWES  
LOUISE RESTRICK

King's College Hospital,  
London SE5 9PJ

MIKE IREDALE

Royal Postgraduate Medical School,  
London W12 0HS

JEFF MEECHAM-JONES

London Chest Hospital,  
London E2 9JX

MARK SANDERSON

Guy's Hospital,  
London SE1 9RT

- Dhillon DP, Winter J, Morgan MDL, Wardlaw AJ, Treacher DF, Woodcock A, et al. Respiratory medicine: fighting for survival. *BMJ* 1992;305:427. (15 August.)

## AIDS and ethics in Birmingham

**EDITOR.**—The article "AIDS and ethics in Birmingham: a betrayal of trust" by Michael Fitzpatrick categorically asserts that a breach of confidentiality was perpetrated by a "member of the Birmingham AIDS team."<sup>1</sup>

There is not a group of individuals so designated, but we believe that the use of this phrase is grossly