Short form 36 questionnaire

Sir—In a recent article by Ziebland,1 the author attempted to assess the usefulness of the short form 36 (SF 36) health status questionnaire in measuring health gain in population surveys. This was done by using recently published normative data from the Oxford lifestyle survey to calculate effect sizes for hypothesized changes in reported health status. The author hypothesized: (i) an improvement in mean reported health status of people in social class V to that of people in social class I, and (ii) an improvement in mean reported health status of 55–65 year olds to that of men and women 10 years their junior. She found that the size of the effect measured by the SF 36 was small to moderate, and concluded from this that the SF 36 was an inadequate tool for assessing population changes in health. This paper provides a classic example of an inappropriately applied statistical method leading a researcher to draw a completely inappropriate conclusion. As I hope to demonstrate, the correct statistical analysis applied to the same data using the same hypothesis leads to the opposite conclusion—that is, that the SF 36 is extremely sensitive to small changes in a population’s perceived health status, even using samples of around 3–400, and is therefore useful in monitoring health gains at population level.

Ziebland is quite correct when she describes the effect size as “a measure of the relative magnitude of a change”, and quotes Cohen’s rule of thumb of 0.2, 0.4, and 0.8 as indicating small, moderate, and large clinical effects respectively.2 However, to say that an effect size is moderate and that the magnitude of a change is not the same as saying that it is a measure of an instrument’s responsiveness. This is because the effect size is determined by four factors: the actual change in health status occurring in the population, the sample size, the variation in health in the population, and the responsiveness of the measure. The reason Ziebland found small to moderate effect sizes in her paper is obvious from the data presented in her tables; the differences in perceived health status between social classes, and between 55–65 year olds and those 10 year younger are only small to moderate. For example, in her paper Ziebland reports a small effect size of 0.23 for the hypothetical change in mean general health (as measured by the SF 36) for men and women of social class V to that of social class I. This is not surprising given the small mean difference (4.8%) between these groups.

A much more helpful method of assessing the usefulness of the SF 36 in measuring health gain in population surveys is to use confidence interval analysis.3 The table presents 95 and 99% confidence intervals for the hypothesized change in mean health status, as measured by the SF 36, for men and women in social class V to that of social class I, using the data reported in Ziebland’s paper. The results indicate quite clearly that even with the small to moderate differences in perceived health status (as measured by the SF 36) observed between social classes V and I, and with samples of 230–380 people in each group, there is less than 1% probability that these observed differences occurred by chance.

The conclusion therefore is that the SF 36 is responsive enough to detect quite small changes in health in a population. Whether or not these small changes are “clinically” or “socially” significant improvements is a matter for health professionals, politicians, and the public to decide.

DANNY RUTA
Toynbee Health Board, PO Box 75, Dundee


Reply

Danny Ruta is concerned that it is an inadequate grasp of statistics which has led me to conclude that community surveys using the SF 36 would lack the sensitivity to assess the impact of health policies. While I am quite happy to agree that the effect size may not be the only statistic appropriate for assessing changes within a population, it has been identified as one of the most often employed for this sort of analysis.1 It is also irresistible (because the error is presumed by a touch of pomposity) to point out that he is mistaken in asserting that sample size is among the ingredients of the effect size calculation, which includes only the mean and standard deviation of the scores.

The hypothetical changes in reported health status were selected as deliberately extreme examples: whether small simple mean differences which range from 4.8% to 9.6% should be considered small, moderate, or dramatic is not something which can be considered in abstract.4 Certainly, the ability of an instrument to detect a difference between the health status of social classes I and V cannot be taken as the final arbiter of responsiveness. It is therefore with some surprise that I see such effort being taken to demonstrate that these (observationally quite uncontroversial) differences would occur by chance less often than 1% of the time in samples of 230–380 per group. This is hardly the issue since policies which could be expected to have even a tenth of the impact of such a change are extremely rare. My contention is that if even such gross changes would elude detection by the approach which has been most frequently used for this purpose, then perhaps other methods of identifying community health gain, such as focused studies using non-survey methods,1 should be considered.

SUE ZIEBLAND
General Practice Research Group, Gibson Building, Radcliffe Infirmary, Oxford OX2 6HE


Smoking and health promotion in Nazi Germany

Sir—Hermann Brenner’s letter seems to consider that our article “Smoking and health promotion in Nazi Germany” should have contained a “careful review of more informative individual-level studies” of interventions aimed at reducing smoking. This seems to rather spectacularly miss the central point of our piece, which is that to understand smoking behaviour in populations, some knowledge of the historical and social background is required. By discussing the possible reasons for the continuing high levels of smoking in Germany, backed up by a cohort analysis stretching back to those who initiated their smoking during or before the second world war, and not referring to the possible long-term influence of one of the most dramatic (and fortunately, in what it accompanied, historically unique) prohibitionist movements the world has seen, seems bizarre. This is especially the case when the reason Brenner cites for the remaining high rates of smoking—the lack of restrictive smoking policies in workplaces and on transport, together with a paucity of health education activity among youngsters— are exactly those which the Nazis implemented, with little success.

Understanding behaviours as complex as smoking requires a considerably more sophisticated view of how the world is than one which sees individual-level motivation as primary.5 The biological and psychological dependence induced by sustained nicotine use and the no less fundamental economic dependence of capitalism on expanding potential markets cannot be considered separately.6 The social function of smoking and the contributions of the status as a smoker to the construction of identities must also be considered.7 Of relevance to our article, the possible conscious adoption of smoking as a form of resistance cannot be easily dismissed.8 Thus smoking among adolescent women can simultaneously represent a transgression against gender stereotyping. In the Nazi period, the Hitler Youth surveillance teams, at least, understood this point. In a Hitler Youth report on the anti-authoritarian Hammer Youth, the contribution of smoking to general degeneracy was clear: “The dance music was all English and American. Only swing dancing and jitterbugging took place. At the entrance to the hall stood a notice on which the words “Swing prohibited” had been altered to “Swing requested”. . . . The dancers were an appalling sight. None of the couples danced normally, there was only swing of the worst sort. Sometimes two boys danced with one girl; sometimes several couples formed a circle, linking arms and jumping, slapping hands, even rubbing the back of their heads together, and then, bent double, with the top half of the body hanging loosely down, their hair flapping into the space, they dragged themselves round practically on their knees. When the band played a rumba, the dancers went into wild ecstasy . . . The band played wilder and
of health, such tools do not attain the degree of rigour demanded of a precise scientific instrument. More worryingly, the text could also be criticised from the very perspective that it intends to promote. It could be argued that a fundamental tension exists between the relatively enlightened thoughts on concepts of health contained within the introductory chapter and the highly formal and quantified tools that subsequently follow. In reviewing discrepancies between the SF-36 measure of general health and associated qualitative assessments of health, Hill, Harries, and Popay have already suggested that irrespective of any intended sensitivity, quantitative tools can never match the subtlety gained from more open ended methods. This tension is left largely unresolved in the text.

Bowling herself points out that it is easy to criticise established tools. It is also arguable that in a policy context which demands quantifiable data to inform decision making, such tools are practical, pragmatic, and inherently useful. If this is the case, Bowling’s book provides an important step towards the development of measures of quality of life that are both sensitive and rigorous.

SANDY WHITELAW
Lecturer in Health Promotion, University of Manchester


BOOKS RECEIVED


NOTICES

International Congress on Women, Work and Health. Barcelona, 18/20 April 1996. For further information, please contact: Dona, Salut i Qualitat de Vida, O/Parris 150 1º, 2 Barcelona 08036, Spain. Tel/fax: 322 65 54.

16th International Congress of Nutrition. Montreal, QC, Canada, 27 July-1 August 1997. For further information, please contact: Congress Secretariat, IUNS ‘97, National Research Council Canada. Building M-19, Montreal Road, Ottawa, ON, Canada K1A 0R6. Tel: (613) 993-7271. Fax: (613) 993-7250.

16th Annual New England Epidemiology Summer Program. 10 June-5 July 1996, 5 and 10 day courses at Tufts University in Medford/Boston, Massachusetts. For more information contact: The New England Epidemiology Institute, Dept PA-JECH, One Newton Executive Park, Newton Lower Falls, MA 02162-1450. Tel: (617) 244-9669; e-mail: epidemiol@aol.com.

Corrigendum

Smoking and health promotion in Nazi Germany

The letter which appeared in October 1995 (49(5):555) was submitted by George Davey Smith and Matthias Egger. The letter is reprinted on p 109.

Society for Social Medicine Abstracts

The abstracts which appeared in the October 1995 issue (vol 49(5):534-54) were from the September 1995 scientific meeting of the Society for Social Medicine, not the 1994 meeting.