Author's reply: We greatly appreciate Professor and Dr Holland's generous comments on our study and would hasten to reassure them that being unorthodox was never our intention. In our paper the relationship between sudden death and prior psychiatric contact is expressed both as a relative risk and as a standard mortality ratio with results further analysed by diagnostic grouping. The design of this study involved starting with a register of all those reported to the coroner as sudden death and then tracing their prior psychiatric histories. It may well have been more credible to have started from the psychiatric case register and to have traced rates of sudden death in patient cohorts, but that was not the methodology employed. We are in the process, however, of analysing just such a study focusing on a large cohort of patients who have been treated for schizophrenia, and we hope to present this for publication in the near future. This forthcoming paper will, we hope, better address the quite proper concerns raised by these correspondents.

P. E. Mullen Monash University and Victorian Institute of Forensic Mental Health, Department of Psychological Medicine, PO Box 266, Rosanna, Victoria 3084, Australia

Analysis of drop-out data in treatment trials

Sir: Wearden *et al* (1998) report the results from a randomised, double-blind, placebo-controlled treatment trial of fluoxetine and graded exercise for chronic fatigue syndrome. In the statistical analysis of their data they use last observation carried forward (LOCF) to replace values missing because subjects dropped-out of the study, and use change scores to estimate treatment effects. Similar analyses are to be found in many other recent papers in the *Journal*, although it is now well-known that both are poor compared with alternatives. LOCF involves highly unlikely assumptions, for example, that the (unobserved) post-drop-out responses remain frozen at their last observed value, and change score estimators have greater variance than estimators based on analysis of covariance – consequently using change scores is less powerful. Full explanations are available in Frisson & Pocock (1992), Everitt (1995, 1998) and, in particular, chapter 7 of Senn (1997).

Everitt, B. S. (1995) The analysis of repeated measures: a practical review with examples. *Statistician*, **44**, 113–135.

— **(1998)** Analysis of longitudinal data: Beyond MANOVA. *British Journal of Psychiatry*, **172**, 7–10.

Frisson, L. & Pocock, S. J. (1992) Repeated measures in clinical trials: analysis using mean summary statistics and its implication for design. *Statistics in Medicine*, **11**, 1685–1704.

Senn, S. (1997) *Statistical Issues in Drug Development*. Chichester: Wiley.

Wearden, A. J., Morriss, R. K., Mullis, R., *et al* (1998) Randomised double-blind, placebo-controlled trial of fluoxetine and graded exercise for chronic fatigue syndrome. *British Journal of Psychiatry*, **172**, 485–490.

B. S. Everitt Department of Biostatistics and Computing, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF

Validity of the Chinese version of the Edinburgh Postnatal Depression Scale

Sir: Lee *et al* (1998) purport to demonstrate the validity of the Edinburgh Postnatal Depression Scale (EPDS) in a population of Chinese women. We would interpret the study as demonstrating the very poor properties of the EPDS – at least in this Chinese version.

If we assume that major depression criteria are taken as the 'gold standard' against which the EPDS is being validated, that the point prevalence of major depression is approximately 5% and, for a cut-off value of 9/10, that the sensitivity is 82% and the specificity 86%, then if 1000 women are screened for major depression, then 174 cases would be identified but of these, 133 would be false positives. It would be tendentious to assert that an instrument has good operating properties if it is wrong three times more often than it is right.

The problem, which is a generic one with screening tools, is that when the base rate of the phenomenon in question is low, the specificity of the tool must be very high in order to avoid the number of false positives outweighing the number of true positives. Lee *et al*'s paper demonstrates exactly this point. Contrary to the authors' conclusions, their data indicate the poor utility of the EPDS for population screening and this is clearly of some concern given that the scale is commonly used for this purpose in clinical practice.

Lee, D. T. S., Yip, S. K., Chiu, H. F. K., *et al* (1998) Detecting postnatal depression in Chinese women. Validation of the Chinese version of the Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry*, **172**, 433–437.

C. Hawley, T. Gale Hertfordshire Neurosciences Research Group, East Hertfordshire NHS Trust, Mental Health Unit, QEII Hospital, Howlands, Welwyn Garden City, Herts AL7 4HL

Author's reply: Drs Hawley and Gale base their calculation on the assumption that major depression is the gold standard against which the EPDS is being validated. This assumption, however, is untrue.

The EPDS is designed to screen for both major and minor depression in recently delivered women. In the original validation study by Cox (1987), as well as in subsequent validations in other cultures, the EPDS was validated using major depression plus minor depression as the gold standard criterion. We followed the same methodology in our study.

The point prevalence of major depression plus minor depression in our study subjects is approximately 12%. If the calculation of psychometric properties is based on this prevalence, the positive predictive value will be 44%, as we reported. A calculation based on 5% prevalence will underestimate the positive predictive value of the rating scale.

Cox, J. L., Holden, J. M. & Sagovsky, R. (1987) Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry*, **150**, 782–786.

D. T. S. Lee Department of Psychiatry, The Chinese University of Hong Kong, Shatin NT, Hong Kong

Compliance, adherence, concordance . . .

Sir: Contributors to the *Journal* (Healey *et al*, 1998; Kemp *et al*, 1998) are still using the term 'compliance'. Psychiatrists, of all specialities, should move away from the attitude that patients should be compliant (i.e. submissive, yielding, acquiescent, obedient). The time has come to recognise that compliance is an outdated concept. Contributors to the *British Medical Journal* (e.g. Lewin, *et al*, 1998) have recently used the term 'adherence' (state of adhering, steady attachment).

Furthermore, the term 'non-compliance' indicates that either the patient is completely compliant or the opposite. However, adherence is better conceptualised along a continuum, with the poles being 100% adherence (extremely rare) and no adherence (extremely rare), with a majority falling somewhere in between.

Healey, A., Knapp, M., Astin, J., *et al* (1998) Cost-effectiveness evaluation of compliance therapy for people with psychosis. *British Journal of Psychiatry*, **172**, 420–424.