

PATTIE, A. H. & GILLEARD, C. S. (1979) *Manual of the Clifton Assessment Procedures for the Elderly* (CAPEBERS). Sevenoaks: Hodder & Stoughton.

#### Asian patients and the HAD scale

SIR: Surely Chaturvedi is rather too severe in his criticism (*Journal*, January 1990, 156, 133) of Nayani's translation of the HAD scale into Urdu (*Journal*, October 1989, 155, 545–547)? He may be right that some researchers: "... have the impression that mere translation of an instrument is sufficient to make it applicable for use in populations of different ethnic or linguistic backgrounds", but a careful reading of Dr Nayani's report shows that he is not one of them. His research intention was not to *use* the scale to measure an Asian population but to *determine its usefulness* (i.e. to find out to what extent it could be applicable and what changes it might need), addressing, in fact, those very issues which Dr Chaturvedi thought were being ignored. Moreover, it is clear that Dr Nayani understands the limitations of word-for-word translation. Various authorities, he says, "... have emphasised the importance of translation of the concept rather than the literal translation of sentences. The HAD was translated on this principle. ...".

Dr Chaturvedi has opened a can of worms. Probably everyone would agree that if we take rating scales that are validated in one culture only, and use them in other cultures without modification, we can obtain nice neat columns of figures which don't mean anything. On the other hand, if we use different measuring scales, each one culturally appropriate and valid in its place of origin, the results will be more ethnographically satisfying and probably more clinically useful. The snag is that we won't be able to use those results for inter-group or international comparisons; and epidemiology is important.

How can we escape from this dilemma? The usual compromise seems to be to start with a well-known rating scale and translate it, then twist and bend it a bit, knocking off a few apparent irrelevances and substituting one or two 'cultural' features, and hope for the best. Is this right? Is there a better way? If compromises are in order, are there some general rules or principles? How many changes can be made to a rating scale before it becomes a different scale? Any? Of course, a scale taken out of its context should be revalidated; but what does that mean – recalibration against a local clinically-selected reference sample, or something more than that? Are there differences (in this respect) between instruments which identify diagnostic categories, and

instruments used only within an agreed category to quantify severity or measure change over time?

If those who are wise in such matters could offer some guidance, I am sure the rest of us would be grateful.

PHILIP RACK

University of Leeds  
Department of Psychiatry  
15 Hyde Terrace  
Leeds LS2 9LT

#### HAD and ROC

SIR: Razavi *et al* (*Journal*, January 1990, 156, 78–93) investigate the characteristics of HAD scale in cancer patients. We have some observations concerning the reporting of such research findings.

Firstly, the HAD scale was devised in order to provide clinicians and researchers with estimates of the presence and severity of two separate emotional disorders: anxiety and depression. It was not devised in order to provide a 'global' concept of the presence of psychiatric disorder as does the General Health Questionnaire. There have been several instances of research reports based upon summation of the two subscale scores of the HAD, but this should not be done. Dr Razavi *et al* later present validation for the two subscales separately, and find the performance of the anxiety scale to be relatively poor; this is to be expected when the gold-standard for HAD *anxiety* is the presence or absence of *depression* (with or without adjustment disorder).

Secondly, the purpose of a receiver operating characteristic (ROC) analysis is to illustrate the relationship between false positives and false negatives at different cut-off points on the scale. ROC analyses are analogous to bar-charts – they should convey information more succinctly than the equivalent table. It is the scale points themselves, not the smoothed-out curve, which the reader wishes to examine, in order to judge relative merits of different cut-offs.

The authors state that "the optimal cut-off for the screening of major depressive disorders seems to be 19". This is incorrect. The purpose of displaying the relationship between true positives and false positives is to allow a *choice* of cut-off. The decision will depend on: (a) the prevalence of the target disorder in the study population; (b) the value and feasibility of intervention with cases identified; and (c) the fate which befalls those patients assigned to the wrong category.

An increasing number of reports of psychometric test data are being presented in terms of ROC analysis. As noted above, one purpose of the ROC chart is