

dysmorphophobics seen in the 1960's. To date the numbers are small, but at least some of the patients who had monosymptomatic delusions 20 years ago have become more floridly ill with time and the diagnosis now is that of a schizophrenic illness.

G. G. HAY

Withington Hospital, West Didsbury,
Manchester M20 8LR

References

- HAY, G. G. (1970) Dymorphophobia. *British Journal of Psychiatry*, **116**, 399–406.
MORSELLI, E. (1886) Sulla dismorfofobia e sulla tafefobia *Boll Accad Med (Genova)* VI, 110–19.

COMMUNITY SCREENING FOR MENTAL ILLNESS

Benjamin, De Calmer and Haran (*Journal*, February 1982, **140**, 174–80) claim to have shown that the General Health Questionnaire (GHQ) is “unsuitable as a screening instrument for mental illness in the community” (p 174). Since the GHQ is one of the most widely used psychiatric screening questionnaires, it is important that their claim be rigorously examined. Two issues arising from their paper require consideration.

1. The respondents were all women aged between 40 and 49 (a demographic subgroup which comprises about 6 per cent of the population). The results cannot therefore be taken to apply to men or to young or old women. Indeed, the sample was not representative of 40–49 year old women, since membership was confined to those of caucasian origin, who “were still able to pass through a ‘natural’ menopause and who could cooperate with multiple investigations of physical, mental and social state”. From a random sample of 228, only 100, or 44 per cent, met these criteria. The sample is thus not representative of anything at all. The authors are, it seems, aware of this, and they defend themselves on p. 179 by observing that “validation studies of the GHQ–30 in a consulting setting do not appear to be affected by demographic variables”. Yet earlier in their paper (p. 174) they observe, as part of their reason for doing the study, that “it is questionable whether responders to such a questionnaire will behave in an identical manner regardless of how they are identified or the circumstances in which the questionnaire is presented”. The importance of the representative nature of a sample (assuming that one wishes, as the authors do, to generalise from it) is a cornerstone of epidemiological investigation, and cannot be dismissed as irrelevant.
2. Even if the sample were representative of the

population (or some definable subgroup), there remains the issue of whether Benjamin *et al*, have in fact shown the GHQ to be “unsuitable as a screening instrument”. Their principal reason for drawing this conclusion is the low sensitivity ($18/33 = 55$ per cent) found when GHQ was compared with the Clinical Interview Schedule (Table III, p. 176), primarily because the questionnaire tended to miss chronic cases. It is worth noting that the number of cases on which this finding is based (33) is relatively small, so that the 95 per cent confidence limits, 38 per cent and 72 per cent, are widely spaced.

Apart from this, sensitivity and specificity are measures of the *validity* of a questionnaire, which is not at all the same as its potential as a screening instrument. To assess this, the relevant indices are the *predictive values* (Galen and Gambino 1975). The positive predictive value (PPV) is the probability that a screened positive will be a “true case”: for Benjamin *et al*'s, data (Table III), the value is 0.78. The negative predictive value (NPV) is the probability that a screened negative will be normal, also 0.78 for the present data. Thus, 8 out of 10 high scorers will be cases, and 8 out of 10 low scorers will not be.

It may well be said that this is unacceptable. Whether, however, the modified GHQ (15 items, Likert scoring) does any better depends on the purpose to which the screening exercise is to be put. It can be calculated from Table VI of Benjamin *et al*, (inspection of which shows that the modified GHQ fails to identify 12 per cent of cases, and not 4 per cent as they claim) that the positive and negative predictive values are 54 per cent and 94 per cent respectively. Thus, if not missing cases were the prime consideration, then the modified version is better; if it were more important to identify *only* cases, then the original version is better, since only about half of the screened positives are cases when the modified version is used. In any case, it is highly unlikely that a questionnaire derived on so atypical a sample will have general validity.

It appears that further studies are required before the claim that “the GHQ is inappropriate as a screening instrument for mental illness in the community” can be substantiated.

PAUL WILLIAMS

*Institute of Psychiatry, Denmark Hill,
London SE5 8AH*

References

- GALEN, P. S. & GAMBINO, S. R. (1975) *Beyond Normality: the Predictive Value and Efficiency of Medical Diagnosis*. New York: Wiley.