Regular Review

Identifying psychiatric illness among general medical patients

DAVID GOLDBERG

The recognition of psychiatric illness by physicians and surgeons comes about either by the doctor noticing some cue that suggests psychological disorder or because the patient's complaints cannot be accounted for by known organic disease. In either case the doctor must be able to follow up with appropriate questions which will enable him to make a confident diagnosis of a psychiatric disorder.

That sounds straightforward, but it is not. Over half of all medical patients with psychiatric illnesses diagnosable according to research criteria will not have their illnesses detected by the medical staff looking after them[1-4]—for five main reasons. Firstly, many such patients do not provide any cues, either non-verbal or verbal, that suggest a psychological disorder, though they will readily describe their symptoms if they are asked directly. Secondly, patients often mention depression or anxiety at the beginning of the interview together with their presenting somatic symptoms, yet only the latter are picked up and discussed further. Thirdly, medical histories are often taken in conditions of little privacy: patients may admit problems in a side room that were denied on the open ward.[1] Fourthly, a known organic cause for a patient's symptoms does not exclude a psychiatric disorder; indeed, such disorders will often be found to be responsible for exacerbating the patient's complaint. Finally, and rather sadly, even when they suspect a psychiatric disorder many clinicians are not confident of their ability to make psychiatric assessments.

One approach to the problem is for the clinician to make use of the bewildering variety of pencil and paper tests that may be given to patients in order to detect those who are likely to have a psychiatric disorder. Most such illnesses turn out to be anxiety states and depressive illnesses, and the screening tests are mainly concerned with the symptoms of these. None of them make diagnoses, but several will provide a profile of scores in addition to a general score indicating the probability that a disorder is present. The account which follows is mainly concerned with those scales which have been well validated in general hospital settings. Since symptoms related to anxiety are correlated with symptoms related to depression many clinicians prefer a single scale that gives an indication of anxiety-depression.

General screening tests

The most widely used screening test is the general health questionnaire (GHQ), which is available in versions as short as 12 items and as long as 60.[5] Over 50 well conducted validity studies have been published, of which seven were in medical outpatients or inpatients. The GHQ 12 takes only two minutes to complete, while the GHQ 60 takes between 10 and 12 minutes. The best discrimination between cases and non-cases requires adjustment of the threshold score: for example, in general practice settings the best threshold for the GHQ 28 is 5/6, whereas for neurological inpatients it is 11/12, probably because the latter patients are more likely to have many somatic symptoms and social dysfunction from their neurological disease. If the questionnaire is to be used in routine clinical work then the threshold is that point where the probability of "caseness" is 50%; scores above threshold represent an increasing probability of the patient being a "case." A small validation study should be carried out if the questionnaire is to be used for research, and the validity coefficients used to compute the predicted prevalence of illness in the entire population screened[5]—the percentage of patients with high scores must not be supposed to be the same as the probable prevalence of illness.

The symptom rating questionnaire (SRQ) as a 24 item screening questionnaire developed by the World Health Organisation for use in developing countries in general medical settings.[6] A recent comparison between the GHQ 12 and the symptom rating questionnaire in Brazil showed them to be about equally effective.[7]

Scaled screening tests

Several scaled screening tests—detecting specific symptoms—are available, and most can also be used as general case detectors by adding the scaled scores together. The symptom check list (SCL) is available in a variety of lengths, the SCL 90 having nine subscales, and the SCL 25 only two, depression and anxiety. A popular version has three additional scales—somatisation, obsessive-compulsive, and interpersonal sensitivity. These scales have been validated in general medical settings in the United States: the main problem proved to be a lower positive predictive value due to more false positive results (see table).[8,9]

The scaled general health questionnaire or GHQ 28 consists of four subscales for somatic symptoms, anxiety and insomnia, social dysfunction, and severe depression. It is at least as good as the GHQ 30 as a case detector, despite the fact that the latter questionnaire has had items excluded that are likely to be positive in physical illness. The Rotterdam scales[10] were designed for use with patients with cancer, but there are only preliminary data on their validity.[11] The hospital anxiety depression scale (HAD), unlike the others, has scales derived from clinical experience rather than factor analysis: it consists of two sets of seven questions with four point response scales.[12] (The validity coefficients quoted in

the table were calculated by counting "doubtful cases" as normal.) Each scale has been shown to correlate well with independent ratings of depression and anxiety. The Middlesex Hospital questionnaire is a self administered inventory consisting of six scales: anxiety, phobic anxiety, obsessional, somatic, depressive, and hysterical. It has not been validated in the same way as the other inventories, however, and cannot therefore be directly compared with them for use in general medical settings.[12a]

Validity coefficients for psychiatric screening tests

	References	Specificity†	Sensitivity‡	Positive predictive value (30% prev)§
General screening tests:				
GHQ 60	2, 5, 21, 23	73-93	67-88	57-75
GHQ 30	5, 22, 23	74-86	72-100	55-75
GHQ 12*	7	79	83	64
SRQ*	6, 7	72-85	73-83	64
Scaled inventories:				
SCL	8, 9	71-72	73-84	52-56
GHQ 28	1, 2, 3, 25	74-92	72-92	54-82
HAD (depn)	12	94	67	70
HAD (anx)	12	76	87	61
Rotterdam	10, 11	76	80	59
Depression screening tests:				
BDI* (short)	15	82	86	67
CES-D*	20, 26	83-94	60-64	60-72
Leeds*	18	85	97	73
SADD*	19	90	86	79

*Study not done in a general medical setting.
†Specificity = true negatives as percentage of non-cases.
‡Sensitivity = true positives as percentage of cases.
§Positive predictive value = probability of high scorer turning out to be a case, at 30% prevalence.
Abbreviations are spelt out in text of article.

Screening tests for depression

The short form of the Beck depression inventory (BDI) consists of 13 questions with four point responses. It correlates very well with the long form of the Beck depression inventory,[13 14] and validity coefficients in the +0·75 range have been reported with the Hamilton rating scale for depression among medical inpatients.[4] This latter scale has been extensively used and has high internal consistency and reliability. The only study which allows validity coefficients shown in the table to be calculated was in general practice.[15] The Center for Epidemiologic Studies depression scale (CES-D) consists of 20 items rated on four point scales.[16] The questions were drawn from impeccable sources, and the scale has been validated against research diagnostic criteria in several large community studies, but it has not been validated on medical inpatients, and both its sensitivity and its positive predictive value have been disappointing in community settings. (The positive predictive value is the probability that someone with a high score will turn out to have the disease at subsequent clinical examination. For any screening test the positive predictive value rises with prevalence, so that the low values reported for this scale in community settings merely reflect the low prevalence of depression in random samples of the population: the values shown in the table look much better because they have been calculated for 30% prevalence.) The Leeds depression scale consists of only six questions, which are scored with four point scales.[17 18] The validity coefficients shown in the table are not really comparable with the other studies, since the investigators knew whether respondents were "patients" or "normals," rather than carrying out blind assessments on a sample of medical patients. The screening test for the World Health Organisation's standardised assessment of depressive disorders (SADD) is not really a screening test in the conventional sense at all but merely a check list of exclusion criteria followed by eight symptoms, of which the patient must have at least two.[19] This screening test was designed for use by psychiatrists, and would probably not be of much use to physicians.

Positive predictive value

For the practising clinician the most important thing is the likelihood that someone with a high score will turn out to be a "case." One screening test may be compared with another if all are standardised to the same prevalence. In the table we have used a prevalence of 30%—an average in most general medical settings. The figures apply to the generality of patients with scores above the threshold: for an individual patient the higher the score above the threshold the higher the positive predictive value. Finally, the various validity studies cannot be compared with one another in a precise way, since they were carried out using different criteria for case identification and in very different cultural settings. Nevertheless, the broadly similar validity coefficients in very different places suggest an impressive consistency in the way people become psychologically unwell: the same psychological and psychophysiological symptoms define minor illness across the world.

DAVID GOLDBERG

Professor of Psychiatry,
University of Manchester,
Manchester M20 8LR

1 Bridges K, Goldberg D. Psychiatric illness in patients with neurological disorders. *Br Med J* 1984; **289**:656-8.
2 Maguire GP, Julier D, Hawton K, Bancroft J. Psychiatric morbidity and referral on two medical wards. *Br Med J* 1974;i:268-71.
3 Moffic H, Paykel ES. Depression in medical in-patients. *Br J Psychiatry* 1975;**226**:346-53.
4 Schwab JJ, Bialow MR, Clemmons RS, Holzer CE. Hamilton rating scale for depression with medical in-patients. *Br J Psychiatry* 1967;**113**:83-8.
5 Goldberg DP. *Manual of the general health questionnaire*. Windsor: NFER Publishing Co, 1979.
6 Harding TW, Arango MV, Baltazar J. Mental disorders in primary health care. *Psychol Med* 1980;**10**:231-41.
7 Mari JJ, Williams P. A comparison of the validity of two psychiatric screening questionnaires in Brazil. *Psychol Med* (in press).
8 Glass RM, Allan AT, Uhlenhuth EH. Psychiatric screening in a medical clinic. *Arch Gen Psychiatry* 1978;**35**:1189-95.
9 Goldberg DP, Rickels K, Downing R, Hesbacher P. A comparison of two psychiatric screening tests. *Br J Psychiatry* 1976;**129**:61-7.
10 De Haes JC. Klachtenlijst voorkanker patienten: erste ervaringen. *Nederlands Tijdschrift voor de Psychologie* 1983;**38**:403-22.
11 Trew M, Maguire GP. *A further comparison of two screening instruments*. Amsterdam: EORTC, 1982. (Study group monograph.)
12 Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;**67**:361-70.
12a Crisp AH, Jones MG, Slater P. The Middlesex Hospital questionnaire: a validity study. *Br J Psychol* 1978;**51**:269-80.
13 Beck AT, Rial WY, Rickels K. Short form of depression inventory: cross validation. *Psychol Rep* 1974;**34**:1184-6.
14 Beck AT, Beck R. Screening depressed patients in family practice. *Postgrad Med* 1972;**52**:81-5.
15 Salkind MR. Beck depression inventory in general practice. *J R Coll Gen Pract* 1969;**18**:267-71.
16 Radloff LS. The CES-D scale. A self-report depression scale for research in the general population. *Applied Psychological Measurement* 1977;**1**:385-401.
17 Kearns NP, Cruickshank KJ, McGuigan KJ, Riley SA, Shaw SP, Snaith RP. A comparison of depression rating scales. *Br J Psychiatry* 1982;**141**:45-9.
18 Snaith RP, Bridge GWK, Hamilton M. The Leeds scales for the self-assessment of anxiety and depression. *Br J Psychiatry* 1976;**128**:156-65.
19 Sartorius N, Davidian H, Ernberg G. *Depressive disorders in different cultures*. Geneva: World Health Organisation, 1983.
20 Boyd JH, Weissman MM, Thompson WD, Myers JK. Screening for depression in a community sample. *Arch Gen Psychiatry* 1982;**39**:1195-200.
21 Byrne P. Psychiatric morbidity in a gynaecological clinic: an epidemiological survey. *Br J Psychiatry* 1984;**144**:28-34.
22 Kumar R, Robson KM. A prospective study of emotional disorders in childbearing women. *Br J Psychiatry* 1984;**144**:35-47.
23 Medina-Mora ME, Padilla GP, Campillo-Serrano M. The factor structure of the GHQ: a scaled version for a hospital general practice service in Mexico. *Psychol Med* 1983;**13**:355-61.
24 Myers JK, Weissman MM. Use of a self-report symptom scale to detect depression in a community sample. *Am J Psychiatry* 1980;**137**:1081-4.
25 Rabins PV, Brooks BR. Emotional disturbance in multiple sclerosis patients: validity of the GHQ. *Psychol Med* 1981;**11**:425-7.
26 Roberts RE, Vernon SW. The CES-D scale: its use in a community sample. *Am J Psychiatry* 1983;**140**:41-5.