

Detecting anxiety and depression in general medical settings

D Goldberg, K Bridges, P Duncan-Jones, D Grayson

Abstract

To aid general practitioners and other non-psychiatrists in the better recognition of mental illness short scales measuring anxiety and depression were derived by latent trait analysis from a standardised psychiatric research interview. Designed to be used by non-psychiatrists, they provide dimensional measures of the severity of each disorder. The full set of nine questions need be administered only if there are positive answers to the first four. When assessed against the full set of 60 questions contained in the psychiatric assessment schedule they had a specificity of 91% and a sensitivity of 86%.

The scales would be used by non-psychiatrists in clinical investigations and possibly also by medical students to familiarise them with the common forms of psychiatric illness, which are often unrecognised in general medical settings.

Introduction

Psychiatric disorders in general medical patients often go undetected.^{1,7} Recent research into ways of training general practitioners in the better recognition of psychiatric illness^{8,9} has been based on full assessments, which have usually entailed administering long research interviews. It is unrealistic to expect clinicians to ask so many questions, many of which may be unnecessary.

Latent trait analysis, developed by Rasch,¹⁰ offers the potential for refining instruments to detect psychiatric disorders in general medical settings.¹¹ This method of analysis provides an intuitively acceptable picture of the relation between symptoms and underlying disease processes. It is possible to calculate two characteristics about each symptom measuring a particular dimension, the "threshold" and the "slope." The threshold of a symptom is that point on the latent trait where probability of having the symptom is 0.5 and is thus a measure of severity of the symptom. It follows that symptoms with higher thresholds are less common than those with low thresholds, so that arranging the symptoms in order of increasing threshold reflects their increasing importance as indicators of the latent trait. The slope measures how well the symptom discriminates between subjects at such a threshold: a symptom with a high slope will have a great change in frequency of occurrence for a fairly small change in the value of the underlying trait.¹²

A short interview to be used by non-psychiatrists should consist of a small set of screening questions to establish whether a disorder is likely to be present and a set of further "probe" questions, which are asked only if positive replies are obtained to the screening questions. Latent trait analysis is ideal in the development of such an interview as the screening questions can be chosen to relate to symptoms with low thresholds but good slopes and the probe questions to relate to those with high thresholds and adequate slopes.

In a recent latent trait analysis applied to data obtained from patients attending 15 general practitioners in Manchester we found that anxiety and depression accounted for all the common relations between their psychiatric symptoms.^{6,12} A score could be calculated for the position of each patient on the latent traits for anxiety and depression.¹² Our purpose was to derive a short interview to be used by general practitioners and physicians that would give results

almost as good as those obtained by a psychiatrist using a standardised research interview.

Methods

The data set was drawn from a total of 427 patients. Those with a low probability of disturbance (those who had a low score on the general health questionnaire and were rated as normal by the general practitioner) were undersampled. The research psychiatrist (KB) carried out 283 interviews with the psychiatric assessment schedule, which consists of 60 questions. He made various diagnoses based on the third revision of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-III)*¹⁴ of the American Psychiatric Association in 144 patients but made no diagnosis in 139 cases. The 139 patients without diagnoses were weighted for a sample size of 283 in order to re-create the original 427 patients.

The first step was to choose screening questions for anxiety and depression. This was done by selecting for each latent trait the four symptoms that had the lowest thresholds combined with satisfactory slopes.¹² In selecting the probe questions we omitted symptoms that depended on the clinical experience of the psychiatrist (such as observed anxiety and observed depression) together with symptoms that are often due to organic disease, such as loss of appetite. Five probe questions about symptoms with high thresholds and reasonable slopes were selected for each latent trait. Table I shows the screening and probe questions and their slopes and thresholds.

TABLE I—Scales for anxiety and depression showing discriminatory ability of each symptom on its appropriate latent trait (slope) and severity of each symptom (threshold)¹²

	Slope	Threshold
<i>Anxiety scale</i>		
Core symptoms:		
Keyed up, on edge	3.88	0.37
Worrying a lot	9.20	0.38
Irritability	2.19	0.49
Difficulty relaxing	1.78	0.75
Supplementary symptoms:		
Poor sleep	1.65	0.76
Headaches, neck aches	1.65	0.86
Autonomic symptoms	1.53	1.03
Health worries	1.10	1.24
Delayed sleep	1.34	1.25
<i>Depression scales</i>		
Core symptoms:		
Low energy	1.18	1.09
Loss of interest	2.65	1.23
Loss of confidence	1.68	1.60
Hopelessness	1.50	1.74
Supplementary symptoms:		
Inefficient thinking	1.06	1.92
Poor appetite—weight loss	0.76	2.43
Early waking	0.88	2.48
Felt slowed up	0.90	3.18
Felt worse in mornings	0.44	5.23

The next step was to devise cut off thresholds for each scale that would allow the clinician to ask the additional probe questions of as few patients as possible. On each scale this was done by cross tabulating scores on the screening questions with scores on the complete set of questions measuring the latent trait that we used in the Manchester study.^{6,12} Each possible threshold score was considered in terms of advantages (saved effort) and disadvantages (missing high scorers and administering probe questions to those who would turn out to have low total scores on the scale). It was

Department of Psychiatry,
University of Manchester,
Manchester M20 8LR
David Goldberg, DM, professor
of psychiatry
Keith Bridges, MRCPsych,
research fellow

NH and MRC Social
Psychiatry Research Unit,
Canberra, Australia
Paul Duncan-Jones, PHD,
mathematical psychologist
(Dr Duncan-Jones has since
died)
David Grayson, PHD,
mathematician

Correspondence to: Professor
Goldberg.

then necessary to decide the cut off score on each of the new scales which best corresponded to the standards used by current research diagnoses. Generalised anxiety disorder was used as a criterion diagnosis for the anxiety scale, and major depressive disorder for the depression scale. Optimal thresholds were computed for each scale by trading off sensitivity against specificity.

The final step was to prepare a scale for use by non-psychiatrists by describing symptoms in simpler, less technical language than in a research interview. Thus "subjective inefficient thinking" became "poor concentration," and "autonomic anxiety" was converted to a short checklist of common autonomic symptoms.

For each scale we computed two scores. The scale threshold, calculated by adding the positive responses to the core questions, determined whether the patient was to be asked the probe questions. The cut off score, calculated by adding the positive responses on the entire scale, indicated whether the patient was likely to have a clinically important disturbance of anxiety or depression.

Results

As all the patients had been asked all the questions on the scales thresholds could be identified for the number of positive responses to the screening questions to determine whether or not the patient needed to be asked the remaining questions on the scale. It proved possible to lower the threshold to two positive answers to the first four questions on the anxiety scale and only one on the depression scale, so that 265 (62%) of the patients in the present sample did not need to be asked the supplementary questions on anxiety and 311 (72%) did not need to be asked the supplementary questions on depression. The advantage was 35% less work when administering the anxiety scale and 40% less work for the depression scale. Less than 1% of patients who were diagnosed as either anxious or depressed by the entire scale would have been missed if the supplementary questions had not been asked. A more serious disadvantage was that of wasted effort: at the scale thresholds adopted 38 (23%) patients who were asked the supplementary questions on anxiety and 39 (34%) who were asked the supplementary questions on depression did not have high scores on the entire scales. With these thresholds for supplementary questions a mean of 5.9 questions were asked on the anxiety scale and 5.3 on the depression scale.

Table II shows the scores of patients with generalised anxiety disorder, major depressive disorder, and a heterogeneous group of other diagnoses made with DSM-III criteria. Well over 80% of patients with anxiety or depressive disorders had high scores on the appropriate scale, and few cases of psychiatric illness were missed by the new scales. The conventional validity coefficients for the new scales compared with diagnoses made from research interviews using DSM-III criteria were calculated. The anxiety scale had a sensitivity of 82% and a positive predictive value (any DSM-III diagnosis at a prevalence of 33%) of 0.56; the depression scale had a sensitivity of 85% and a positive

TABLE II—Number (percentage) of diagnoses made with anxiety and depression scales compared with diagnoses made with DSM-III criteria

DSM-III diagnosis	Score of >5 on anxiety scale	Scores of >2 on depression scale	Low scores on both scales
Generalised anxiety disorder (n=67)	55 (82)	16 (24)	9 (13)
Major depressive episode (n=55)	7 (13)	47 (85)	1 (2)
Other (n=22)	10 (45)	6 (27)	9 (41)
None (n=284)	16 (6)	12 (4)	259 (91)

predictive value of 0.85. The overall specificity (the percentage of patients without psychiatric disorders who scored low on both scales) was 91%; the overall sensitivity (the percentage of patients diagnosed as having a psychiatric disorder scoring above the threshold on at least one scale) was 86%.

The box shows the form of the two scales as they

Anxiety scale

(Score one point for each "Yes")

- 1 Have you felt keyed up, on edge?
- 2 Have you been worrying a lot?
- 3 Have you been irritable?
- 4 Have you had difficulty relaxing?

(If "Yes" to two of the above, go on to ask:)

- 5 Have you been sleeping poorly?
- 6 Have you had headaches or neck aches?
- 7 Have you had any of the following:
trembling, tingling, dizzy spells, sweating,
frequency, diarrhoea?
- 8 Have you been worried about your health?
- 9 Have you had difficulty falling asleep?

Depression scale

(Score one point for each "Yes")

- 1 Have you had low energy?
- 2 Have you had loss of interests?
- 3 Have you lost confidence in yourself?
- 4 Have you felt hopeless?

(If "Yes" to ANY question, go on to ask:)

- 5 Have you had difficulty concentrating?
- 6 Have you lost weight (due to poor appetite)?
- 7 Have you been waking early?
- 8 Have you felt slowed up?
- 9 Have you tended to feel worse in the mornings?

Interpretation:

Add anxiety score, add depression score. Patients with anxiety scores of five or depression scores of two have a 50% chance of having a clinically important disturbance; above these scores the probability rises sharply.

might be used by doctors. A patient with a score at the cut off for either scale (that is, five symptoms of anxiety or two symptoms of depression) has a 50% chance of having a clinically important disturbance, and above these scores the probability rises sharply.

Discussion

Researchers may still wish to administer pencil and paper tests^{15 16} to assess anxiety and depression in patients, and those with sufficient time and training will want to use standardised research interviews to diagnose common mental disorders.¹⁷⁻²⁰ The scales presented here are designed for use at the bedside by clinicians and have been designed so that lengthy scales do not need to be administered to patients with a low probability of having an affective disorder.

Though the scales should not supplant the usual history taking by an experienced doctor so that discussions between the doctor and patient become stereotyped and automatic, they have two important uses. In practice, doctors often rely on a few (sometimes oddly chosen) questions to gain an impression of the degree of depression or anxiety of patients, and the content of the new scales may encourage them to broaden their repertoire of questions. Some guidance on the severity of disturbance represented by each symptom is provided by the measures given in table I.

The scales would be useful in surveys requiring measures of the severity of psychological disturbance in groups of patients if no psychiatrist were available to

administer standardised research interviews. Scores can be used directly as dimensional measures to examine correlations with other such measures. Another application would be to use the proportion of patients with scores above the cut off scores as an indicator of the prevalence of affective illness in a particular population. The scales could also be administered by medical students to patients on the general wards of hospitals, to make the students more aware of the forms of affective illness common in this setting.

As these scales have been derived from questions administered by a psychiatrist during a research interview they cannot be assumed to perform equally well when administered by non-psychiatrists. Further studies in which patients in a general medical setting are interviewed on separate occasions by a non-psychiatrist using the new scales and by a psychiatrist using a research interview will confirm the validity of the scales in non-specialist use.

USE OF THE SCALES IN CLINICAL PRACTICE

The box shows the scales as they might be used by clinicians. Our original interview asked about symptoms experienced in the past month; this period seems realistic for patients in general medical settings. Symptoms that are no longer present, or that are present only in mild degree, should be ignored. Symptoms other than those mentioned here can be asked about, but the score should be obtained from only the symptoms mentioned in the scales. Each question is given the same score, but the items are given in order of increasing severity, so that symptoms further down each scale are encountered in patients with more severe disorders.

The cut off scores given here have been determined in general practice settings; possibly some upward revision may be necessary to obtain the best results

among patients on medical or surgical wards. The scales can now be used in clinical settings and compared with results of diagnosis with more time consuming methods.

- 1 Goldberg DP, Blackwell B. Psychiatric illness in general practice: a detailed study using a new method of case detection. *Br Med J* 1970;iii:439-43.
- 2 Mofitt H, Paykel ES. Depression in medical in-patients. *Br J Psychiatry* 1985;126:346-53.
- 3 Skuse D, Williams P. Screening for psychiatric disorder. *Psychol Med* 1984;14:365-78.
- 4 Bridges K, Goldberg DP. Psychiatric illness in patients with neurological disorders: patients' views on discussion of emotional problems with neurologists. *Br Med J* 1984;289:656-8.
- 5 Sireling K, Paykel ES, Freeling P. Depression in general practice. *Br J Psychiatry* 1985;141:113-9.
- 6 Goldberg DP, Bridges K. Screening for psychiatric illness in general practice: the general practitioner versus the screening questionnaire. *J R Coll Gen Pract* 1987;37:15-8.
- 7 Brodaty H, Andrews G, Kehoe L. Psychiatric illness in general practice: why is it missed? *Aust Fam Physician* 1982;11:625-31.
- 8 Goldberg DP, Steele J, Smith C, Spivey L. Training doctors to recognise psychiatric illness with increased accuracy. *Lancet* 1980;ii:521-4.
- 9 Gask L, Goldberg DP, Lesser A, Millar T. Improving the psychiatric skills of the general practice trainee: an evaluation of a group training course. *Med Educ* 1988;22:132-8.
- 10 Rasch G. *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute of Educational Research, 1960.
- 11 Duncan-Jones P, Grayson D, Moran P. The utility of latent trait models in psychiatry epidemiology. *Psychol Med* 1986;16:391-405.
- 12 Goldberg DP, Bridges K, Duncan-Jones P, Grayson D. Dimensions of neuroses seen in primary care settings. *Psychol Med* 1987;17:461-70.
- 13 Dean C, Surtees P, Sashidharan S. Comparison of research diagnostic systems in an Edinburgh community sample. *Br J Psychiatry* 1983;142:247-56.
- 14 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 3rd ed. Washington, DC: APA, 1980.
- 15 Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;67:361-70.
- 16 Goldberg DP, Hillier V. A scaled version of the general health questionnaire. *Psychol Med* 1978;9:139-45.
- 17 Hamilton M. The assessment of anxiety states by rating. *Br J Med Psychol* 1959;32:50-8.
- 18 Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56-62.
- 19 Goldberg DP, Cooper AB, Eastwood MR, Kedward HB, Shepherd M. A standardised interview suitable for use in community surveys. *British Journal of Preventive and Social Medicine* 1970;24:18-24.
- 20 Wing JK, Cooper JE, Sartorius N. *The measurement and classification of psychiatric symptoms*. London: Cambridge University Press, 1974.

(Accepted 8 June 1988)

Increased rates of previous hysterectomy and gynaecological operations in women with osteoarthritis

T D Spector, G C Brown, A J Silman

Department of Clinical Epidemiology, The London Hospital Medical College, London E1 2AD
T D Spector, MRCP, Wellcome research fellow
A J Silman, FFCM, senior lecturer

Department of Rheumatology, St Bartholomew's Hospital, London EC1
G C Brown, BSC, medical student

Correspondence to: Dr Spector.

Osteoarthritis is more common in women around the age of 50 than in men of the same age.¹ We investigated in a case-control study possible gynaecological or hormonal influences that might account for this observation.

Patients, methods, and results

The cases comprised women aged 35-70 with osteoarthritis attending one of six local outpatient clinics. Two groups of similar ages were used for comparison: 314 women with rheumatoid arthritis attending the same clinics and a group drawn from a random sample of 1000 women obtained from electoral registers in Greater London. Subsequently we also studied 346 women from a local general practice. Each woman received a postal questionnaire that contained questions on menarche, menopause, parity, use of the contraceptive pill, hysterectomy, and date of onset of disease when applicable. Diagnoses and gynaecological details in the hospital groups were verified by inspection of medical records. The response rates were 83% (327/393) in the osteoarthritis group, 91% (286/316) in the rheumatoid arthritis group, and 70% (243 replies in the

age range) in women from the electoral register. Owing to inadequate data the final sample sizes were those shown in the table.

The table shows the main results. Women in the osteoarthritis group reported a higher rate of hysterectomy (28%) than those in the rheumatoid arthritis group (14%) and population control groups (15% and 11%). The median time from hysterectomy to diagnosis was six years in the osteoarthritis group and four years in the rheumatoid arthritis group. Further gynaecological information was available for most of these women. Oophorectomy was reported by 24 out of 48 (50%) women with osteoarthritis and 10 out of 25 (40%) with rheumatoid arthritis. The commonest surgical indications were menorrhagia or fibroids or both (45 out of 57 (79%) with osteoarthritis and 18 out of 29 (62%) with rheumatoid arthritis). There was also a higher rate of dilatation and curettage (82 (42%) v 52 (26%)) and other gynaecological problems (110 (52%) v 75 (34%)) among the women with osteoarthritis.

The timing of hysterectomy in relation to the onset of disease was examined in the two arthritis groups. The crude relative risk of hysterectomy at least one year before referral to hospital for osteoarthritis was 2.8 (95% confidence interval 1.7 to 4.6). Using age at first symptoms did not alter the estimates. The relative risk adjusted for age was 2.9 (1.8 to 4.7); adjustment for other variables did not appreciably alter the estimate.

Comment

The increased rate of hysterectomy and other gynaecological operations in women with osteoarthritis