British Journal of Psychiatry, (1990), **157**, 679–686

The Interval General Health Questionnaire

P. G. SURTEES and P. McC. MILLER

An adaptation of the LIFE procedure was used with a self-report questionnaire to assess 173 medical students, close to their enrolment and again about six months later. The assessment procedure enabled changes in symptoms of anxiety and depression over the interval between interviews to be recorded and the subsequent classification of the symptom patterns into course categories. While women scored markedly higher than men at initial assessment, this was not so at follow-up. Almost 25% of students reported the co-occurrence of at least five psychological symptoms at some time during the six months. For many students these were of a persistent nature. The adapted GHQ meets the dual demands of restricted interview time and the need to assess the more minor psychological conditions.

One of the enduring problems confronting psychiatry concerns how best to measure the very phenomena upon which the discipline is based, and in recent years this measurement problem has become a dominant feature of the research literature. Central to this effort is the fundamental and reasonable concern with applying universal standards to provide a rational framework for the continued evolution of the assessment and classification of psychiatric phenomena. Of course it will always be the case that particular assessment techniques may be best suited to particular settings and that developments in psychiatric nosology may have only tangential or, at best, uncertain relevance to large bodies of activity in psychological research. An alternative but equally legitimate strategy of inquiry to that involving the use of diagnostic concepts, has relied upon those assessment methods whose primary objective is to provide a simple enumeration of symptoms at a particular time or during a particular period. Such methods usually are based on the self-report of symptoms by questionnaire and thereby carry with them certain hazards concerning their meaning and interpretation. Nevertheless, they are widely used and have been found to be of particular value, especially for screening purposes, in identifying individuals with an increased likelihood of a current psychiatric disorder.

Both standardised psychiatric interviews and self-report assessments, in different ways, offer complementary approaches which are broadly adequate to meet the range of measurement requirements concerning 'current' psychiatric status. A continuing challenge is to provide viable solutions to the problem of assessing *change* in psychiatric status over time, be it with reference to a diagnostic scheme or for more minor, perhaps rather transient changes in symptom state. The development of standardised psychiatric interviews reflects the continuing evolution of psychiatric diagnostic schemes, with new criteria being embodied in the interviews as they are agreed upon. In general, while there has been a preoccupation with the development and imposition of standardised diagnostic criteria, there has been a relative neglect of measures specifically designed to document the *course* of episodes of psychiatric disorder. One outstanding exception has been the work originally initiated by Dr Robert Shapiro and continued by Dr Martin Keller. The interviewing technique they developed, the Longitudinal Interval Follow-up Evaluation (LIFE), was designed for a multicentre study of depressive illness (Keller *et al*, 1987). A special feature of this approach is that it assesses variation over time of the 'psychiatric status' of the interviewee in accord with the particular diagnostic criteria and scheme chosen for study. The method has facilitated a detailed examination of, for instance, factors associated with the recovery and relapse of major depressive disorder (e.g. Keller *et al*, 1982*a*,*b*) and with the risk of chronicity (Keller *et al*, 1986).

The present paper reports on the application of some of the principles of the LIFE method to the assessment of the course and change in minor spells of psychiatric distress and thereby seeks to help bridge the divide between advances made in formal psychiatric interview assessments and self-report methods. Through this combination of methods, it is also hoped to underpin the advantages already demonstrated for the criterion as opposed to the conventional aggregate scoring procedure of a symptom check-list (e.g. see Uhlenhuth *et al*, 1983; Murphy *et al*, 1985, 1987; Murphy, 1986).

As part of a study designed to investigate the interrelationship between health, coping resources, social support, social adversity and personality in newly enrolled medical students, an opportunity arose to

https://doi.org/10.1192/bjp.157.5.679 Published online by Cambridge University Press

evaluate a method of documenting change over time in self-reported symptoms. This part of the study had the following specific aims:

(a) to assess the prevalence of psychiatric symptoms in newly enrolled medical students at entry and over the ensuing six months of their training, with a special emphasis being given to the examination of sex differences
(b) to classify students into categories on the basis of the patterns of reported symptoms over time and to investigate whether there is a sex-specific differential adaptation to entry to medical school
(c) to evaluate an aggregate versus a criterion-based approach to summarising change in psychiatric symptoms over time
(d) to determine the six-month inception rate of spells of psychiatric symptoms and to identify whether there are sex differences in these rates.

Method

Following pilot work and interviewer training, all first-year medical students at the University of Edinburgh were approached for interview within four weeks of their enrolment. The interview included sections dealing with the following areas: demographic factors, health, social contacts, coping resources, and personality. All students seen at the initial interview were approached approximately six months later for a follow-up interview. This paper is concerned with only one component of the assessment, psychological state.

As part of the assessment, the 30-item General Health Questionnaire (GHQ; Goldberg, 1978; Goldberg & Williams, 1988) was administered, both at the initial and follow-up interviews. Since its development, the GHQ has been used in a variety of forms (e.g. from 12–60 items) to facilitate the detection of psychiatric disorder in a wide variety of settings and cultures (e.g. Tarnopolsky *et al*, 1979; Banks, 1983; Chan & Chan, 1983; Mari & Williams, 1985; Bridges & Goldberg, 1986; Williams *et al*, 1986; Boardman, 1987). Its choice for this study ensured comparison in some respects with other work, but the strategy to be described below in relation to particular symptoms is generally applicable and therefore need not be tied or limited to the GHQ.

The principles of the LIFE method depend in part upon determining the times of change in psychiatric status ratings (PSRs) within a study interval. The PSRs relate directly to the extent to which diagnostically significant symptoms are present and co-occur. Ratings made at one interview are 'drawn forward' at follow-up, based upon new information gathered at the follow-up interview. The interviewer seeks to identify only those times which correspond to *changes* in PSRs. For present purposes, this principle was adopted though limited to certain symptoms selected from the GHQ. Our primary objective therefore was to be able to document the changes in those symptoms over the study period.

The process of selecting items from the GHQ-30 was guided not only by the results of published papers but also by our desire to impose a criterion-based scoring method. Two subsets of six items each were selected from the 30-item GHQ that were considered to be broadly related to depressive and anxiety conditions. Guidance in selecting items was obtained from the manual of the GHQ (in particular results concerning the factor structure of the questionnaire) and from those symptoms that are commonly included within diagnostic criteria for depressive and anxiety disorders and for which the GHQ-30 contained appropriate questions. Several papers provided results which influenced item choice.

Goldberg & Williams (1988) have reported the results of a number of principal-component analyses of the GHQ-30 and GHQ-12. Between 48% and 71% of the total variance was accounted for, in the case of the GHQ-30, by from four to ten significant components. Of these studies, one undertaken by Goldberg *et al* (1976) on 1310 United States general practice patients accounted for 51.5% of the total variance using the 30-item GHQ. The first of the four significant factors in the varimax rotation accounted for 20.6% of the variance and was labelled 'depression and anxiety'. Subsequently, Goldberg & Hillier (1979) published a scaled version of the GHQ based upon a varimax rotation of the 60-item GHQ, 48% of the variance being accounted for by four scales each consisting of seven questions; of these scales, scale B was labelled 'anxiety and insomnia', and scale D 'severe depression'. The six chosen 'anxiety' items were in close accord with those items comprising scale B in the 28-item GHQ (scale B item 2 being taken to be approximately equivalent to item 3 on the GHQ-30).

Only four of the items chosen to represent our 'depression set' formed part of scale D of the 28-item GHQ (items 24, 25, 29 and 30). For reasons similar to those guiding choice of the anxiety-related items, two additional depression-related items were also included (item 22 "Been feeling unhappy and depressed?" and item 23 "Been losing confidence in yourself?" – items G1, sensitivity 86%, and G2, sensitivity 78%, respectively in the original factor analyses).

In addition to the above findings, the results obtained from the latent-trait analysis of symptom data (whether obtained from formal psychiatric assessments (see Goldberg *et al*, 1987, Table 4) or based on the GHQ (see Duncan-Jones *et al*, 1986, Table 2) provided additional support for item choice. It should also be noted that the overlap between our final choice of items and the 12-item GHQ was limited to only five questions.

The two sets of selected items were cast into a computer-generated pro-forma that included the GHQ item number, the wording of each chosen item, the individual student scores for the items at initial interview, the date of initial interview and columns to allow the follow-up scores to be recorded (see Appendix). Upon completion of the follow-up GHQ, the interviewer was required to record the scores of each of the selected items on the personalised pro-forma and then follow the interview procedure for completion of the ratings covering the follow-up interval. Symptom levels at the two interviews provided anchor points to facilitate establishing each student's symptom status over the study interval. Students were not directly questioned about each

https://doi.org/10.1192/bjp.157.5.679 Published online by Cambridge University Press

TABLE I
GHQ and I-GHQ scores at initial and follow-up interviews, by sex

	Initial interview						*Follow-up interview*					
	All (n = 175)		*Males only (n = 96)*		*Females only (n = 79)*		*All (n = 173)*		*Males only (n = 95)*		*Females only (n = 78)*	
Scale scores: mean (s.d.)												
GHQ	5.93	(5.85)	4.58	(4.69)	7.56*	(6.69)	4.70	(5.89)	4.56	(5.41)	4.87	(6.45)
I-GHQ:												
all items	2.44	(2.49)	1.80	(1.93)	3.22*	(2.85)	1.76	(2.43)	1.72	(2.18)	1.81	(2.72)
anxiety set	1.69	(1.49)	1.38	(1.36)	2.08*	(1.57)	1.19	(1.56)	1.21	(1.47)	1.17	(1.67)
depression set	0.75	(1.22)	0.43	(0.79)	1.14*	(1.51)	0.57	(1.05)	0.51	(0.93)	0.64	(1.18)
Percentage scoring at or above cut-off point												
GHQ: % scoring ≥5	49.7		40.6		60.8		35.8		34.7		37.2	
I-GHQ all items, % scoring ≥2	53.7		41.7		68.4		37.6		40.0		34.6	
anxiety set, % scoring ≥2	47.4		36.5		60.8		32.4		35.8		28.2	
depression set, % scoring >0	38.3		28.1		50.6		32.9		31.6		34.6	
'Sensitivity'[1]	90.8		87.2		93.8		88.7		87.9		89.7	
'Specificity'[1]	83.0		89.5		71.0		91.0		85.5		98.0	

1. Relationship between I-GHQ (all items, cut-off point ≥2) and GHQ (usual cut-off point of ≥5) as criterion.
*$P<0.001$.

month of the study period in order to establish symptom ratings – questions were directed only towards identifying those times when changes in the selected symptoms had occurred which corresponded to the GHQ scale points.

Two scoring methods were adopted for this study. In the first, symptoms were scored according to the conventional binary scoring (0 or 1) of GHQ items; in the second, a criterion-based approach was adopted. In order to apply the second method, two items were endowed with greater significance than others to reflect the importance usually placed upon them in diagnostic schemes. The item "Been feeling nervous and strung-up all the time" (item 28 GHQ-30; item B7 scaled GHQ) was chosen as the 'pivotal' item for the 'anxiety set' because it appeared to be a reasonable proxy for the key questions used to establish generalised anxiety disorder by formal psychiatric interviews imposing the Research Diagnostic Criteria (RDC; Spitzer *et al*, 1978) or DSM–III–R criteria (American Psychiatric Association, 1987). This particular item (under conventional GHQ scoring) appears to have acted as an effective discriminator of the 'severely ill group' in the original GHQ item analysis (see Goldberg & Williams, 1988; Appendix 1, item G11, item sensitivity 72%). Similar reasons influenced the choice of item 22, "Been feeling unhappy and depressed?", selected as the key criterion item for the set of depression-related symptoms.

An example of a completed form, together with the associated interviewer guide notes, is given in the Appendix. This interval form of the questionnaire will be referred to as the 'I-GHQ'.

Results

Of the 181 students eligible for inclusion, interviews were successfully obtained from 175 (96.7%) within four weeks of their enrolment. Five to six months following the initial interview all students were approached once again for a follow-up assessment, and 173 interviews (98.9%) were completed. All of the students were single and they ranged in age from 17 to 25 years with only three aged over 20 years.

The results shown in Table I were obtained using the conventional binary scoring (0 or 1) of all the 30 GHQ items and a similar method of scoring the 12 items of the I-GHQ. While at initial interview the female students reported symptom levels significantly higher ($P<0.001$) than the males on all measures shown in Table I, the difference was no longer present at follow-up. Further indications of the extent to which this student sample scored highly on the full questionnaire are given by Fig. 1. The pattern of results revealed by the GHQ was repeated for the I-GHQ.

Also in Table I, the conventional cut-off point applied to the 30-item GHQ is contrasted with others based on the I-GHQ. At initial interview almost half the students scored at or above the conventional cut-off points, while at follow-up about 35% were still at or above this level. Estimates of 'sensitivity' (in this instance the proportion of 'GHQ-30 cases' correctly identified) and 'specificity' (the proportion of 'GHQ-30 normals' correctly identified) associated with the chosen cut-off point on the I-GHQ are also shown in Table I to indicate the extent to which the same students are being identified by both forms of the questionnaire.

https://doi.org/10.1192/bjp.157.5.679 Published online by Cambridge University Press

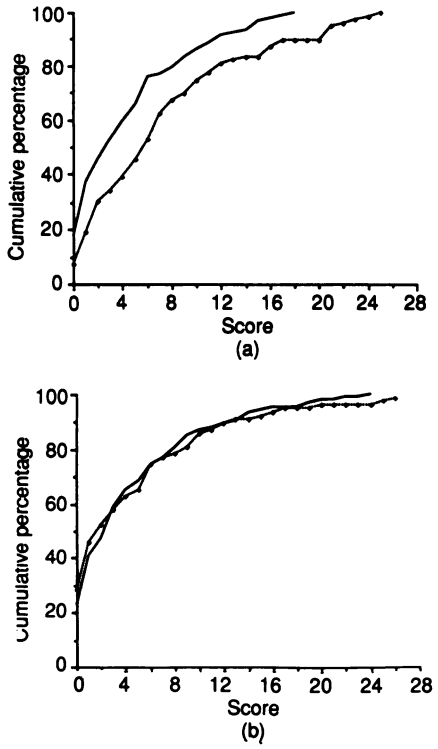


FIG. 1. Cumulative distribution by sex of scores on the 30-item GHQ at both (a) initial and (b) follow-up interviews (——males, $n = 96$; ------females, $n = 79$). A similar pattern was shown with the 12-item I-GHQ.

The results show considerable similarity – perhaps surprising given the brevity of the I-GHQ.

Students were also classified according to the criterion-based approach. In view of the empirical nature of the exercise, several grades of criterion severity were established. The initial (general) criterion adopted was the presence of either (or both) of the two 'pivotal' symptoms, "Nervous and strung-up all the time" and "Feeling unhappy and depressed", with or without any additional symptoms of the ten remaining (all rated according to the conventional GHQ binary scoring): this was considered sufficient within any of the interval monthly time frames to signal a "spell of psychological symptoms" at that time. Subsequently, three criterion levels were imposed where specific numbers of additional symptoms (similarly scored) had to be present in addition to depression and/or anxiety. The first of these involved none or one additional symptom (out of the 10 possible), the second two or three and the third four or more additional symptoms.

The first analysis therefore focused upon the initial (general) criterion mentioned above and (with adjustment when necessary for variation in individual student follow-up times) classified the pattern of ratings according to the categories shown in Table II. For example, a student who at initial interview did not satisfy the minimum criterion but at follow-up reported the necessary symptoms to be present and to have commenced within the study period, would be allocated to the 'new and on-going' subgroup. The category 'multiple' refers to those students who reported several discrete spells of symptoms with each spell meeting the necessary general criterion and being separated by times when the pivotal symptoms were not present. While there was an overall sex difference in the distribution of patterns ($\chi^2 = 12.48$, d.f. $= 5$, $P = 0.029$), it is clear that those patterns potentially of greatest concern were of broadly equal prevalence among the sexes. Similar percentages of male and female students were found to have experienced anxiety and/or depression continuously for most of the follow-up (20.0% of the males and 23.1% of the females).

The degree to which the students fulfilled the more demanding symptom criteria is shown in Table III, which shows that a considerable proportion of this new intake of medical students reported experiencing multiple psychological symptoms at the same time, and that for a sizeable subgroup they were of an enduring nature. Almost a quarter of all students (and a third of female students) had spells of symptoms that included depression and/or anxiety together with four or more additional symptoms, and that lasted for 9.3% of the study time (representing about eight student years). The female students experienced spells of symptoms that were particularly long lasting.

TABLE II
'Spell' course status as determined by the general criterion[1], by sex

'Spell' course status	*All* n	*All* %	*Males only* n	*Males only* %	*Females only* n	*Females only* %	*Male:female* % *ratio*
Well	65	37.6	46	48.4	19	24.4	1.98:1
Recovering	30	17.3	11	11.6	19	24.4	0.48:1
New and recovery	21	12.1	9	9.5	12	15.4	0.62:1
New and on-going	20	11.6	10	10.5	10	12.8	0.82:1
Multiple	22	12.7	12	12.6	10	12.8	0.98:1
Continuous	15	8.7	7	7.4	8	10.3	0.72:1
Totals	173	100	95	100	78	100	

1. Report of either of the two pivotal symptoms, "Nervous and strung-up all the time" or "Feeling unhappy and depressed".

https://doi.org/10.1192/bjp.157.5.679 Published online by Cambridge University Press

TABLE III
Fulfilment of various 'spell' criterion levels over the entire study time, and six-month inception rates

Sample	*Depression and/or anxiety present together with a number of additional symptoms*			
	0 or 1	*2 or 3*	*4 or more*	*Totals*
Overall 'morbidity' levels: %				
All ($n = 173$)				
% meeting each criterion at least once	24.9	38.7	24.9	62.4
mean % of entire study time at each criterion	7.6	13.1	9.3	30.0
Males only ($n = 95$)				
% meeting each criterion at least once	19.0	33.7	17.9	51.6
mean % of entire study time at each criterion	5.0	13.4	5.8	24.2
Females only ($n = 78$)				
% meeting each criterion at least once	32.1	44.9	33.3	75.6
mean % of entire study time at each criterion	10.8	12.7	13.4	36.9
Six-month inception rates: %				
All ($n = 113$)	19.5	23.9	10.6	42.5
Males only ($n = 70$)	14.3	18.6	8.6	34.3
Females only ($n = 43$)	27.9	32.6	14.0	55.8

Percentages are not additive across criterion levels.

As a further indicator of the degree to which these students experienced psychological symptoms, analyses were undertaken to determine the six-month inception rates for each of the criterion-based levels of symptoms. Inception for this study refers to those students who developed spells of symptoms (at each criteria level) for the *first time* during the period of observation. The inception rates therefore exclude those students who at initial interview fulfilled any of the criterion levels being tested. The results are also shown in Table III, for the overall group and separately by sex. It shows for example that, of those 43 female students who were not bothered by either depression and/or anxiety at the time of the initial interview, 14.0% went on to experience symptom spells that not only included depression or anxiety, but four or more symptoms in addition. The inception rate of symptom spells fulfilling this criterion for all students was 10.6 per six months per 100 students.

Discussion

Inevitably many issues remain to be addressed. These concern the selection of items from the GHQ-30 and how they were subsequently analysed, the relationship between the I-GHQ, the GHQ-30 and clinician-based measures of disorder. Questions arise from the substantive findings, in particular how they compare with other populations of a similar age and what may be the advantages of using the I-GHQ beyond one of the more usual forms of the GHQ.

The results of this study have shown that the I-GHQ appears to perform in much the same manner as its 'longer parent' in conventional cross-sectional analyses. A further and much more detailed discussion of the relative performances of the two scales in relation to external clinical psychiatric status forms the basis of the following paper (this issue, pp. 686–693).

This study has clearly indicated the extent to which newly enrolled medical students experience spells of distress during their first few months at medical school. As the students provided the first opportunity to evaluate the I-GHQ, it has not been possible to compare the profile of 'course' categories obtained in this study with any others. However, the GHQ scores at each interview provide one means of comparison. Among 791 British 18–24-year-olds included in a random sample of over 6000 respondents (Cox *et al*, 1987) who completed the GHQ-30, female scores exceeded those of males (females: mean score 4.86, 38% scored ≥ 5; males: mean score 3.53, 30% scored ≥ 5). Comparison of these figures with those shown in Table I reveal that while the scores of the medical students at enrolment were markedly higher than the broadly age-equivalent random sample, at follow-up only the male students still remained relatively high scoring.

A further comparison is with a study by Firth (1986) in which 318 medical students in their fourth year at the universities of Sheffield, Manchester and Leeds completed the 12-item version of the GHQ and were compared with a sample of 222 young employed

https://doi.org/10.1192/bjp.157.5.679 Published online by Cambridge University Press

people. The estimated prevalence of psychological disturbance among the students was 31.2% compared with 9.7% in the young employed. A follow-up of a sample of 170 of the same medical students during their pre-registration year revealed that 50% then had scored four or above on the GHQ-30 (Firth-Cozens, 1987). A study of 523 medical students from the University of Belgrade (Eric *et al*, 1988) assessed both close to enrolment and two years later using the GHQ-60, and supplemented by a psychiatric interview, revealed an overall point-prevalence rate of psychiatric disorder at enrolment of 16.1% (females 21.6%) and two years later 17.5% (females 22.7%).

Evidence of the extent to which medical students in the initial stages of their course meet psychiatric diagnostic criteria has also been demonstrated by Zoccolillo *et al* (1986). In their study of 304 students at medical school in the United States, monthly scores on the Beck Depression Inventory (BDI; Beck *et al*, 1961) were obtained to select students for interview with the NIMH Diagnostic Interview Schedule (DIS; Robins *et al*, 1981). They found a 12% incidence rate (16% of women, 7% of men) of DSM–III major depressive disorder among students during the first two years of medical school together with a significant sex difference in rates. They concluded that depression was at least three times as common among medical students as in the general population.

The medical students in our study were little different from their US counterparts in terms of levels of reported emotional disturbance, almost 25% of the new intake in Edinburgh reporting experiencing the co-occurrence of at least five psychological symptoms during the study period.

Considerable recent research has been directed towards identifying those features of the medical-training environment that are associated with heightened levels of distress in students. Studies designed to look at student factors (e.g. Murphy *et al*, 1984; Notman *et al*, 1984; Lovallo *et al*, 1986; Sobal, 1986; Spiegel *et al*, 1986; Carmel & Bernstein, 1987; Rosenthal *et al*, 1987; Vitaliano *et al*, 1989) have relied mostly on self-report measures with assessments completed at intake and/or during the subsequent years of medical training. It is hoped that the availability of a 'guided' self-report measure of psychiatric symptom course, which is relatively efficient to administer, may facilitate the identification of factors that confer resilience or indicate the vulnerability of students to the largely inevitable stress associated with medical training. Findings concerning relationships between such factors and the changing mental health of the students as measured by the I-GHQ will be addressed in a future publication. Further development of the I-GHQ may involve, for example, giving due allowance in the scoring of symptom items to 'chronicity considerations' (e.g. Goodchild & Duncan-Jones, 1985; Surtees, 1987), and perhaps also to the choice of items and by asking subjects supplementary questions on impairment and help-seeking.

During the last few years there has been an increasing concern with many aspects of medical education, recently culminating in a world conference designed to identify agreed goals and strategies for the future. This conference concluded with a declaration in which specific aspects of medical training including curriculum and examination systems were identified as worthy of improvement (World Federation for Medical Education, 1988). This and other papers have shown that the transition to medical school is highly stressful and it would appear from other published work concerned with later stages of medical training that the educative process continues to be associated with heightened levels of emotional disturbance in the trainees. While it is to be hoped that efforts will be made towards achieving the ideals specified in the declaration, some consideration may also be given to the means of managing the very high levels of stress perhaps sometimes considered to be part of the 'rites of passage' of those undergoing medical training.

Acknowledgements

The authors wish to acknowledge the help of Myrna Kaplan for her assistance in data collection and Deirdre Anderson, Eileen Brand, Katherine Cheshire, Jenny Holden, Gill Ingham, Rita Kirk, Maggie Stevenson and Maura Walker for undertaking the interviews. Invaluable computing assistance was provided by Dilys Rennie in the preparation of interview materials.

References are given at the end of the following paper, pp. 692–693.

Appendix

Interval GHQ–interviewer guide notes

The purpose of this section of the interview is to obtain an overview of certain specific symptoms that each student may have experienced since their last interview. The assessment is based upon items from the General Health Questionnaire (GHQ).

Section format

Two subsets of items have been selected from the GHQ. One set of six items is concerned with anxiety symptoms

https://doi.org/10.1192/bjp.157.5.679 Published online by Cambridge University Press

Scale no.	Item	Score at time 1 (date)	Months: Nov.	Dec.	Jan.	Feb.	Mar.	Score today (date)
	Anxiety set							
2	Lost much sleep over worry	1		1	2			2
3	Having restless and disturbed nights	2						2
14	Felt constantly under strain	1				1	3	4
19	Scared/panicky for no good reason	1		1	2			2
21	Everything getting 'on top' of you	2				2	1	1
28	Nervous and strung-up all the time	2						2
	Depression set							
22	Feeling unhappy and depressed	2	4	3	4	3	2	2
23	Losing confidence in yourself	2	4			4	2	2
24	Thinking of self as a worthless person	2						2
25	Felt that life is entirely hopeless	2						2
29	Felt that life isn't worth living	2	3	2				2
30	Couldn't do anything as nerves too bad	1						1

FIG. 2.

(the anxiety set), the other set of six items with depression related symptoms (the depression set). The 12 GHQ questions form the basis for this section of the interview. A separate interview coding form has been produced (Fig. 2). You will find the following information on each of these forms:

(a) the study code number associated with each student
(b) the two sets of GHQ items, divided according to whether they form part of the anxiety set or the depression set – each item has the scale number associated with the full 30-item questionnaire, as well as the GHQ question
(c) the actual scores for those items recorded by each student at the time of their initial interview is entered in the 'Score at' column: code 1 "Not at all", code 2 "No more than usual", code 3 "Rather more than usual", code 4 "Much more than usual"
(d) Note that the date of first interview also appears in the 'Score at' column
(e) The intervening months for each student are shown at the head of each blank column
(f) Note the 'Score today' column.

Procedure

On completion of the full GHQ, score the 12 items that appear on the form. Attention to the scale numbers on the form will help in searching for the items. Score each item 1, 2, 3 or 4 as above, and enter in the 'Score today' column.

Your aim now is to ascertain the student's symptom status over the interval between the ratings. While space is provided for monthly ratings the student does not have to be asked about how he/she was for each month. Instead probes should be used to determine change points which correspond to the GHQ scaled points. Monthly ratings can then be completed on the basis of identified change points. You should help the student to date change points, by asking such questions as "Was that in November?", "Did that happen before or after Christmas?", etc.

General

Examining the scores at initial interview will indicate symptom status at that time and that can be readily compared with current status for each symptom. Ask probe questions according to the pattern of ratings that each student presents. Do not assume that the interval is symptom free if both interview ratings are 1. In completing the forms, and where no change points are found, simply enter a line between ratings (which will be equal in score value). In general, enter a line covering times of equal ratings.

Fill in today's responses on the questionnaire

Some useful lead questions would be:

"Now I want to look in slightly more detail at some of the items on the General Health Questionnaire."

https://doi.org/10.1192/bjp.157.5.679 Published online by Cambridge University Press

"Last time you were interviewed, that is on (*give date*) you said about "Lost much sleep over worry" that this was (*give student's response*). How have things been since then?"

Now work through each symptom on the form probing to establish change points month by month. Be careful about using today's response, and very careful not to cross-examine the student about inconsistencies. Start each symptom off in the same way.

"When did you begin to feel better (or worse)?"

"Do you think that was a continuation of the way you were feeling at the time of the initial interview or a new period of (difficulty, depression, as appropriate) for you?"

*P. G. Surtees, MSc, PhD, FBPsS; *Member of the Scientific Staff, MRC Biostatistics Unit, 5 Shaftesbury Road, Cambridge CB2 2BW*; P. McC. Miller, BSc, PhD, *Member of the Scientific Staff, MRC Reproductive Biology Unit, Royal Edinburgh Hospital, Morningside Park, Edinburgh EH10 5HF.*

*Correspondence

British Journal of Psychiatry (1990), **157**, 686–693

The Interval General Health Questionnaire and its Relationship with Psychiatric Disorder in Community and General Practice Samples

P. G. SURTEES and M. TANSELLA

This paper contrasts the usual scoring of the GHQ-30 with scoring based upon a criterion approach applied to only 12 of the GHQ items that takes account of the 'diagnostic significance' of the chosen items. Repeat assessments with this shorter form of the questionnaire (the I-GHQ), coupled with procedures to assess change, can be used to provide summary measures of symptom course in those research situations where available interview time precludes the use of more formal assessment methods.

The previous paper (this issue, pp. 679–693) describes a measure and procedure for documenting the presence and course over time of brief spells of anxiety and depression (and associated) symptoms. It explains how 12 items comprising the Interval General Health Questionnaire (I-GHQ) were extracted from the 30-item GHQ. While the selection of the 12 items comprising the I-GHQ was guided by diagnostic considerations (for inclusion with a criterion scoring method), the chosen items had previously been shown to have relatively superior sensitivity in relation to a severely ill group in the development of the GHQ. However, the final choice of the 12 items was empirically based. To underpin the results reported in the initial study, it is necessary to assess the relative performance in case identification of the chosen item subset in contrast to that of the longer form of the questionnaire (the GHQ-30) from which they were drawn.

While a comparison of the scales in one study is informative, the use of data from several study populations would be preferable. Ideally, these studies should be carried out in different settings and also with the application of different psychiatric case criteria. For the present study some attempt has been made to achieve this objective.

This paper is based upon the analysis of three data sets which together go some of the way towards meeting the above requirements for this investigation. As the method and design for each study has been described in some detail already, only an outline is given here.

Method

Study 1

This community study was undertaken in a geographically defined area of north Edinburgh. A random sample of 576 women aged 18–65 years selected from a total eligible population of about 44 000 (1981 census) completed an initial interview, and a subsample of those were re-interviewed six months and one year later. Details of the design and method of the initial survey can be found elsewhere (Dean *et al*, 1983; Surtees *et al*, 1983). Additional information concerning the follow-up interview method is

https://doi.org/10.1192/bjp.157.5.679 Published online by Cambridge University Press