The role of the general health questionnaire in general practice consultations

PATRICIA SMITH

SUMMARY

Background. *The patient self-rating questionnaire is commonly used as a research tool to identify patients with 'unrecognized' depression. There is no evidence to support its use as a clinical tool in general practice.*

Aim. *To determine whether use of the 30-item general health questionnaire (GHQ) is a practical means of increasing identification of 'new' episodes of emotional distress among patients consulting their general practitioner (GP).*

Method. *A randomized controlled trial was carried out in a Scottish new town practice with eight partners. In the waiting room, 1912 patients aged over 14 years and consulting over a 10-month period attempted to complete the GHQ. The 'clinical judgement' group posted the questionnaire into a box then attended the doctor as normal. The 'screened' group presented the questionnaire to the doctor. After the consultation, the doctor completed an assessment questionnaire. The main outcome measures were GHQ scores and doctors' assessments of mental health.*

Results. *In total, 1589 patients were eligible to participate. However, 207 patients in the screened group were excluded because the doctor did not look at the questionnaire. The clinical judgement group (59.7% patients) and the screened group (40.3%) were compared. Although the doctors' diagnoses of distress were low in the clinical judgement group (8.1%), they were significantly greater in the screened group (13.9%) where the diagnosis of depression was doubled. The percentage of patients scoring greater than or equal to 9 (GHQ+) was 21.5% and 21.0% respectively. The level of agreement between the doctors' diagnoses of distress and the questionnaires scoring GHQ+ rose from 19% in the clinical judgement group to 35% in the screened group.*

Conclusions. *The general health questionnaire used in a practice setting increases the identification of patients with emotional distress. However, the use made of the questionnaires in the screened group raises questions of doctor and patient acceptability.*

Keywords: depression, general health questionnaire, randomized controlled trial.

Introduction

EMOTIONAL problems experienced by patients in general practice have been given a high profile in recent years.[1] Up to 20% of the general practice population are estimated to be emotionally disturbed,[2] depression being the commonest diagnosis made under research conditions.[3] Between 70% and 90% of people who develop depression consult their doctor; more so when severely depressed.[4-6] There is also evidence to suggest that treatment of depression in general practice is beneficial to the patient.[7]

P Smith, MBCHB, MRCGP, research fellow, Department of General Practice, 20 West Richmond Street, Edinburgh.

Submitted: 30 April 1997; accepted: 7 April 1998.

Attention has focused on patients with 'unrecognized' depression[8-11] whose condition, despite fulfilling research criteria for a depressive illness, remains undiagnosed when they present at their doctor. The patient self-rating questionnaire is used under research conditions to screen attenders for probable 'psychiatric cases'. The 30-item general health questionnaire (GHQ) was used in this study as an adjunct to the doctors' clinical skills to assist patient assessment. The questionnaire is well validated for use in a general practice setting.[12] It takes only a few minutes for the patient to complete and the score can be easily totalled by the doctor during the consultation.

This study aims to determine the practical application of the GHQ during a routine consultation and to assess its effect on recognition of emotional distress.

Method

This randomized controlled trial was conducted over a 10-month period at a health centre in a Scottish new town. The practice population comprised 10 000 patients, of whom 2097 were approached. The eight general practitioners at the practice, four male and four female, all participated. They each had an equal share of practice workload and a part-time appointment outwith the practice.

Patients were eligible for the study if they were aged 14 years or over, registered at the practice, and attending for a 'non-emergency' appointment. Patients were recruited once only. They were excluded if identified as 'distressed' (as defined by the doctor) at recruitment or during the preceding year: in these circumstances, the doctor will already be sensitized to the diagnosis. The level of distress ranged from trivial and transient to severe and prolonged.

The days of each week were allocated using random number tables to either control or experimental conditions. Under control conditions the patient posted the completed questionnaire into a sealed box before consulting the doctor, creating a 'clinical judgement' group. Under experimental conditions the patient gave the questionnaire to the doctor to scan at the beginning of the consultation forming a 'screened' group.[13]

Doctors conducted one surgery per week under research conditions. They were listed alphabetically, rotating in such a way that, for each doctor, week days were equally represented. Reception staff instructed patients and handed out questionnaires for four days of the week. Surgeries started 10 minutes late to accommodate the study. Immediately after each consultation the doctor completed a questionnaire.

Patient questionnaire

This provided demographic, registration, and attendance details. It also included the 30-item GHQ. A cut-off point of five is used in epidemiological studies;[14,15] at this score there is a 50% likelihood of the patient having a psychiatric condition.[16] However, in a general practice setting, research has shown that increasing the threshold value to nine increases the clinical usefulness of the questionnaire.[11,17,18,19]

The Likert scoring method[12] was explained to the doctors and its interpretation was discussed.

All questionnaires that were attempted were included in the analysis because doctors were likely to respond to a patient scoring nine or more (GHQ+), whether or not the questionnaire was complete. Analysis of scores less than nine (GHQ-) should therefore be interpreted with caution.

Doctor questionnaire

The doctor questionnaire collected the following data:

- Patient factors important to the consultation – physical, psychological, social problems, and other (e.g. repeat prescriptions, sickness certificates).
- Diagnosis of current emotional distress and chronic physical disorders: the classification system of the Royal College of General Practitioners[20] was used.
- Management of patients identified as having emotional distress:
  - 'long' or 'short' discussion about distress as assessed by the doctor,
  - prescription for an antidepressant or tranquillizer, or sick ness certificate for distress,
  - referral to health visitor, community psychiatric nurse, psychologist and psychiatrist,
  - follow-up of the patient – either instructed to make anoth er appointment to discuss distress or expected to initiate further contact.
- Indicators of the practical implications of using the GHQ:
  - duration of consultation (stop clocks were provided),
  - doctor's perceived stress level on a scale of 1 to 7.

Statistical analysis

The data were analysed using SPSS-X. Chi-square tests were used for comparison of groups and a two-tail *t*-test was used for comparison of continuous variables.

Results

Over a 10-month period, 2097 patients were offered a questionnaire (Figure 1). Those patients who refused to participate totalled 185 (8.8%), of which 102 (55.1%) reported forgetting glasses as a reason for not completing it, 13 (7.0%) had language difficulties, 35 (18.9%) had physical problems that precluded writing, and 35 (18.9%) gave no explanation. Because of pre-existing distress, 323 (15.4%) patients were excluded. Of the remaining 1589 patients, 207 (13.0%) were excluded because the doctor, although screening conditions applied, disregarded the questionnaire. Subsequently, 1382 patients were eligible for inclusion: 825 (59.7%) in the clinical judgement group and 557 (40.3%) in the screened group.

The number of patients seen per doctor in the clinical judgement group ranged from 81 (9.8%) to 147 (17.8%) (mean value = 103.1), and in the screened group from 35 (6.3%) to 103 (18.5%) (mean value = 69.6). The wider variation in the screened group is a result of doctors with a high proportion of disregarded questionnaires.

The number of questionnaires with 30 items completed was 747 (90.5%) in the clinical judgement group and 491 (88.2%) in the screened group.

Comparison of clinical judgement groups and screened groups

Patients in the two groups had similar mean age and sex distributions (Table 1).

The percentage of patients scoring GHQ+ was similar in both groups – 21.4% and 21.0%.

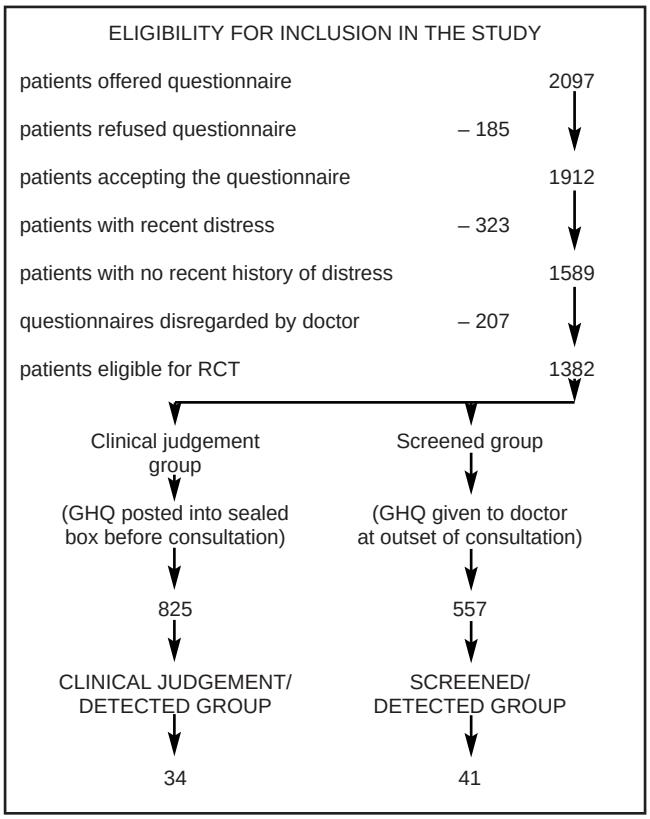

Figure 1. Emotional distress in general practice.

Table 1 shows the influence of the questionnaire by comparing doctor assessment (GP+/-) and questionnaire rating (GHQ+/-) for both groups. Use of the questionnaire in the screened group was associated with a larger percentage of patients being identified as distressed by the doctors. Table 1 also shows that the level of agreement between doctor and questionnaire in identifying distressed patients (GP+GHQ+/GHQ+) had risen from 19.2% in the clinical judgement group to 35% in the screened group. When patients scored GHQ-, the level of disagreement between doctor and questionnaire (GP+GHQ-/GHQ-) was small – 5.1% versus 8.2% respectively.

Comparison of clinical judgement/detected groups and screened/detected groups

Patients with questionnaire scores ≥9, and a doctor diagnosis of distress, formed detected groups (GP+GHQ+).

There were 34 patients in the clinical judgement/detected group and 41 patients in the screened/detected group (Table 2). The sex distribution and mean age in both groups were similar. There were no statistically significant differences between the groups when compared for duration of registration (majority registered for over five years), frequency of attendance at that doctor previously (most attended at least once in the preceding six months), presence of concurrent physical symptoms and chronic health problems. The spread of scores and the mean scores were similar.

Diagnosis

The doctor questionnaire identified 22 different diagnoses of 'new' distress where the consulting doctor was unaware of a pre-

Table 1. Characteristics of clinical judgement and screened groups.

	Clinical judgement group (n = 825)	Screened group (n = 557)	P-value
Sex			0.091[a]
Male	327 (39.6%)	196 (35.2%)	
Female	498 (60.4%)	361 (64.8%)	
Age (years)			0.17[b]
Mean age (SD)	35.8 (15.2)	34.7 (14.3)	
GP GHQ assessment			0.0037[a]
GP+ GHQ+	34 (4.1%)	41 (7.4%)	
GP+ GHQ-	33 (4.0%)	36 (6.5%)	
GP- GHQ+	143 (17.3%)	76 (13.6%)	
GP- GHQ-	615 (74.6%)	404 (72.5%)	
Percentage of agreement between the doctors' diagnosis of distress and the GHQ			
GP+ GHQ+	34 (19.2%)	41 (35.0%)	
GP- GHQ+	143 (80.8%)	76 (65.0%)	
GHQ+	177	117	
GP+ GHQ-	33 (5.1%)	36 (8.2%)	
GP- GHQ-	615 (94.9%)	404 (91.8%)	
GHQ-	648	440	

[a]Chi-square test. [b]Two-tail t-test. SD = standard deviation; GP+ = doctor diagnosis of distress; GHQ+ = score ≥9; therefore, for example, from GP GHQ assessment, GP+ is 34+33 = 67 or 4.1%+4.0% = 8.1% for clinical judgement group and 41+36 = 77 or 7.4%+6.5% = 13.9% for screened group.

Table 2. Characteristics of detected groups.

	Clinical judgement/ detected (n = 34)	Screened/detected (n = 41)	P-value
Sex			0.44[a]
Male	11 (32.4%)	10 (24.4%)	
Female	23 (67.6%)	31 (75.6%)	
Age (years)			0.94[b]
Mean age (SD)	34.8 (14.8)	35.1 (14.6)	
Attended doctor in past six months[c]			0.72[a]
Never	11 (32.4%)	8 (20.0%)	
Once	7 (20.6%)	9 (22.5%)	
2-3 times	8 (23.5%)	14 (35.0%)	
>4 times	8 (23.5%)	9 (22.5%)	
Duration of registration			0.59[a]
>5 years	25 (73.5%)	28 (68.3%)	
GHQ score			0.99[a]
9 to 14	13 (38.2%)	16 (39.0%)	
15 to 20	12 (35.3%)	14 (34.1%)	
21 to 30	9 (26.5%)	11 (26.8%)	
Mean score (SD)	16.4 (6.0)	16.5 (5.0)	0.97[b]
Concurrent medical symptoms	26 (76.5%)	30 (73.2%)	0.74[a]
Chronic health problems	16 (47.1%)	16 (39.0%)	0.78[a]

[a]Chi-square test. [b]Two-tail t-test. [c]missing data = 1. GHQ = general health questionnaire.

Table 3. Management of patients with 'detected' stress.

Doctor management	Clinical judgement/ detected	Screened/ detected	P-values
Long discussion about distress	43.8% (14/32)	26.3% (10/38)	0.13
Short discussion about distress	50% (16/32)	68.4% (26/38)	0.12
Antidepressant prescribed	12.5% (4/32)	2.6% (1/38)	NA
Tranquillizer prescribed	6.3% (2/32)	5.3% (2/38)	NA
Sickline issued for distress	9.4% (3/32)	7.9% (3/38)	NA
Plan to discuss at follow-up appointment			
Doctor initiated	31.3% (10/32)	29.7% (11/37)	0.89
Patient initiated	41.9% (13/31)	68.6% (24/35)	0.030
Referral to secondary services	– (0/32)	7.9% (3/38)	NA

Statistics = chi square; NA = not applicable; where n values differ from Table 2, this is owing to missing data.

vious diagnosis over the past year. A new diagnosis of depression was made in only four (11.8%) patients in the clinical judgement/detected group (endogenous 4) and 10 (24.4%) patients in the screened/detected group (endogenous 4, reactive 4, postnatal 2). A new diagnosis of anxiety state was made in 10 (29.4%) patients in the clinical judgement/detected group and in nine (22.0%) patients in the screened/detected group. Other diagnoses were varied and often non-specific.[21]

Management

Table 3 shows the doctors' management of these distressed patients. There was no significant difference in doctors' perception of discussion about distress as being long or short in duration. Few patients were prescribed antidepressants or tranquillizers in both groups. Less than one in 10 patients were issued with a sick certificate because of distress. Almost one-third of patients were directed by the doctor to make another appointment to discuss their distress. The more vague option, where the patient was expected to initiate a follow-up consultation, was used more frequently, particularly in the screened/detected group. Referral rates were very low in both groups.

Acceptibility of the questionnaire to the doctor

There was wide variation among the doctors in the number of questionnaires that were disregarded in the screened group, from four (2.7%) to 41 (53.9%), with three doctors accounting for 146 (70.5%) of all those disregarded.

Doctor stress

When patients scored GHQ+, three doctors' mean levels of stress were greater in the screened group than the clinical judgement group (Table 4). Two of these doctors disregarded most questionnaires in the screened group. When patients scored GHQ-, one doctor's mean stress levels were significantly increased in the clinical judgement group and one in the screened group.

Table 4. Factors affecting acceptability of GHQ to doctor.

Doctor stress (mean scores by doctor for GHQ+ and GHQ- ratings).

	Clinical judgement group mean stress score (n)	Screened group mean stress score (n)	P-value
GHQ+			
Doctor B	3.1 (8)	4.4 (15)	0.053
Doctor C	1.3 (22)	1.9 (10)	0.043
Doctor F	2.9 (16)	3.7 (15)	0.048
GHQ-			
Doctor A	4.8 (73)	4.4 (59)	0.015
Doctor H	1.7 (111)	2.1 (39)	0.008

Mean consultation length for GP+GHQ+ and GP-GHQ+ groups.

	Clinical judgement group minutes (SD)	Screened group minutes (SD)	P-value
GP+ GHQ+ (detected)	12.8 (7.1)	10.4 (4.7)	0.11
GP- GHQ+	7.4 (3.6)	9.6 (5.6)	0.001

Statistics = two tail t-test; n = number of patients; SD = standard deviation.

Consultation length

There was no statistically significant difference between the detected groups (GP+GHQ+) in the average length of consultation (Table 4). However, when the doctor considered the patient mentally well despite a high score (GP-GHQ+), consultations were significantly longer in the screened group.

Discussion

In this study, as in previous studies based in general practice, the GHQ was associated with a significant increase in the detection of emotional distress.[10,22,23] However, the value of the questionnaire will be influenced by the effect of detection on patient outcome and evidence is, so far, conflicting.[10,24]

Use of the term 'emotional distress' reflects the difficulty in classifying patients who present to the doctor with undifferentiated symptoms. Identification of distress by doctors in the clinical judgement group was low (8.1%; Table 1) in comparison to a similar group in a study from the Netherlands by Ormel *et al*[9] which identified 14% of patients as having a 'psychological disorder' (similar in definition to emotional distress). The diagnosis of depression has important treatment and outcome implications for the patient.[7] The number of patients newly diagnosed as depressed by doctors was low: half the number of cases identified by standardized interview in Ormel's study. Poor diagnostic agreement between doctors for psychiatric disorders[21] and variation in the abilities of doctors to identify depression[25] may explain this. It is also recognized that diagnoses may be reached over a number of consultations.[26] The accuracy of diagnoses made by doctors or questionnaires however, can not be assessed because standardized psychiatric interviews were not conducted.

Acceptability of the questionnaire to the doctor

Although the prospect of having a surgery on a study day was perceived as undesirable, this was in part because the doctors had to complete a questionnaire after each patient. Only one out of eight doctors consulting with patients who scored GHQ- recorded greater stress levels in the screened group compared with the clinical judgement group, which implies that handling the questionnaire itself did not cause undue disruption.

In the screened group, 27.1% of questionnaires were disregarded by the doctor, a higher proportion than expected. This may sometimes have been due to oversight during busy surgeries. However, it seems likely that it was also a means of reducing some doctors' stress levels, and therefore suggests a reluctance to tackle emotional issues at that time. Doctors could be encouraged to make more effective use of the questionnaire and perhaps reduce stress levels by developing an 'in house' protocol for the management of distressed patients. Future studies may more effectively address the problem of disregarded questionnaires.

The doctors disagreed with more questionnaires showing GHQ+ than they agreed with. GP-GHQ+ was associated with a significantly longer average consultation time in the screened group than in the clinical judgement group. Therefore, use of GHQ+ appears to have had some impact within the consultation even though it has not prompted the doctor to make a psychiatric diagnosis. In future, use of the questionnaire may be adapted so that the mental health of patients who score GHQ+ could be assessed during a longer appointment at a future date.

Access to GHQ+ scores in the screened/detected group was associated with identifying more distressed patients but with no significant difference in the level of intervention offered. Therefore, those patients whose distress would otherwise have gone unrecognized were treated similarly to those whose distress was evident to the doctor in the clinical judgement/detected group. Use of the questionnaire does not appear to have major resource implications.

Acceptability of the questionnaire to the patient

Only 8.8% of patients refused to participate in the study. Although reception staff encouraged patients to complete the questionnaire, they did not have the resources to assist them. However, the acceptability of the questionnaire in the longer term may be influenced by the effect on the patient of the doctor not acknowledging a GHQ+ score. In the future, acceptability of self-rating questionnaires could be further investigated by assessing patient satisfaction after the consultation.

Conclusion

In conclusion, the GHQ could be an acceptable addition to the consultation because its use does facilitate the identification of patients who are emotionally distressed. However, the large proportion of questionnaires that were ignored raises concerns both for doctor and for patient acceptability. Also, even when the questionnaires were considered by the doctor, notification of GHQ+ scores was not synonymous with recognition. However, it is important to recognize the longitudinal nature of the diagnostic process in general practice. These findings have implications for modifying future use of the GHQ in a clinical setting.

References

1. Paykal ES, Priest RG. Recognition and management of depression in general practice: consensus statement. *BMJ* 1992; **305:** 1198-1202.
2. Shepherd M, Clare AW. *Psychiatric illness in general practice.* (2nd edition.) Oxford: Medical Publications, 1981.
3. Blacker CVR, Clare AW. The prevalence and treatment of depression in general practice. *Psychopharmacology* 1988; **95:** S14-S17.
4. Brown G, Harris T. *Social origins of depression.* London: Tavistock, 1978.
5. Wing JK, Cooper JE, Sartorius N. *The measurement and classification of psychiatric symptoms.* Cambridge: Cambridge University Press, 1974.
6. Weissman M, Myers JK. Rates and risks of depressive symptoms in

a US urban community. *Acta Psychiatr Scand* 1978; **57:** 219-231.
7. Hollyman JA, Freeling P, Paykel ES, *et al*. Double-blind placebo-controlled trial of amitriptyline among depressed patients in general practice. *J R Coll Gen Pract* 1988; **38:** 393-7.
8. Sireling LI, Paykel ES, Freeling P, *et al.* Depression in general practice: case thresholds and diagnosis. *Br J Psych* 1985; **147:** 113-119.
9. Ormel J, van den Brink W, Koeter MWJ, *et al.* Recognition, management and outcome of psychological disorders in primary care: a naturalistic follow-up study. *Psychol Med* 1990; **20:** 909-923.
10. Johnstone A, Goldberg D. Psychiatric screening in general practice: a controlled trial. *Lancet* 1976; **1:** 605-608.
11. Wright AF, Perini AF. Hidden psychiatric illness: Use of the general health questionnaire in general practice. *J R Coll Gen Pract* 1987; **37:** 164-167.
12. Goldberg D, Williams P. *A users guide to the general health questionnaire.* NFER-Nelson Publishing, 1988.
13. Iliffe S, Mitchley S, Gould M, Haines A. Evaluation of the use of brief screening instruments for dementia, depression and problem drinking among elderly people in general practice. *Br J Gen Pract* 1994; **44:** 503-507.
14. Goldberg DP, Eastwood MR, Kenward HA, Shepherd M. A standardised psychiatric interview for use in community surveys. *B J Preventive and Social Medicine* 1970; **24:** 18-23.
15. Tennant C. The General Health Questionnaire: a valid index of psychological impairment in Australian populations. *Med J Aust* 1977; **2:** 392-394.
16. Goldberg DP. *Manual of the general health questionnaire.* Windsor: NFER, 1978.
17. Finlay-Jones RA, Murphy E. Severity of psychiatric disorder and the 30-item general health questionnaire. *Br J Psychiatry* 1979; **134:** 609-616.
18. Johnstone A, Shepley M. The outcome of hidden neurotic illness treated in general practice. *J R Coll Gen Pract* 1986; **36:** 413-415.
19. Nott PN, Cutts S. Validation of the 30-item GHQ in postpartum women. *Psychol Med* 1982; **12:** 409-413.
20. Royal College of General Practitioners. *Classification of diseases, problems and procedures 1984.* London: RCGP,1984.
21. Jenkins R, Smeeton N, Marinker M, Shepherd M. A study of the classification of mental ill-health in general practice. *Psychol Med* 1985; **15:** 403-409.
22. Zung WWK. The role of rating scales in the identification and management of the depressed patient in the primary care setting. *J Clin Psychiatry* 1990; **51(6):** 72-76.
23. Zung WWK, Macgill M, Moore JT, George DT. Recognition and treatment of depression in a family medicine practice. *J Clin Psychiatry* 1983; **44:** 3-6.
24. Dowrick C, Buchan I. Twelve month outcome of depression in general practice: does detection or disclosure make a difference? *BMJ* 1995; **311:** 1274-1276.
25. Marks JN, Golderg DP, Hillier VF. Determinants of the ability of general practitioners to detect psychiatric illness. *Psychol Med* 1979; **9:** 337-353.
26. Blacker CVR, Clare AW. Depressive disorder in primary care. *Am J Psych* 1987; **150:** 735-751.

Acknowledgements

Thanks are due to the doctors, staff, and patients who participated in this study; and Anne King, Dr Alan Rollo, and the research staff of the Department of General Practice, Edinburgh, for their support, in particular, David Heaney, Jane Hopton, and Michael Porter. This study was funded by the Scottish Home and Health Department.

Address for correspondence

Dr Patricia Smith, Howden Health Centre, Livingston, West Lothian EH54 6TP.