Goldberg's General Health Questionnaire

3 authors, including:

Ulrik F Malt
University of Oslo
443 PUBLICATIONS   **11,864** CITATIONS

The validity of the General Health Questionnaire in a sample of accidentally injured adults

U. F. Malt

Department of Psychosomatic and Behavioural
Medicine, The National Hospital, University of Oslo
(Head: UF Malt) and Department of Surgery, The
Central Hospital of Akershus (Head: N Helsingen)

ABSTRACT – The 20-item version of the General Health Questionnaire (GHQ) was
given to 110 accidentally injured adults during the hospital stay. The patients were
asked to rate their mental state prior to the accident. They also filled in the GHQ
twice during a follow-up period of 28 months. All patients were twice examined by a
psychiatrist who was blind to the patients' GHQ response. The overall values for
sensitivity and specificity were good during the follow-up period and only slightly
better using Chronic scoring procedure (80%–80%). The optimal cutting point for
case screening varied across the different time-periods ($^2/_3$ pre-accident, $^3/_4$ follow-
up). Patients with permanently higher case-scores at follow-up than at pre-accident
were among those most severely affected by the injuries from a psychosocial point of
view. The nine subjects who scored as a case at each of the three points of time all
qualified for a DSM-III diagnosis. Sixty-three percent of the patients with a case
score on two occasions qualified for a DSM-III diagnosis. The study indicates that
GHQ-20 would be a useful measure of the psychosocial state of trauma patients seen
in the course of rehabilitation.

Introduction

Psychological distress and psychopathology are
related to the occurrence of accidents and the
severity of the sustained injuries (1, 2). Psy-
chological distress and psychopathology also
have crucial influence on the short and long-
term adaptation to accidental injuries. Acci-
dental injuries may also cause negative long-
term psychosocial problems (3, 4). According-
ly, the detection of psychological distress and
psychopathology by the surgeon is important
in order to establish optimal treatment condi-
tions and facilitate the process of rehabil-
itation.

The General Health Questionnaire (GHQ) (5)
and the John Hopkins Symptom Distress Check-
list (SCL) (6) are the two most used screening
instruments for distress and psychopathology.
Both questionnaires exist in long (example
GHQ-60; SCL-90) and short ended versions (i.e.
GHQ-12 & GHQ-20; SCL-22) and have been
found to be reliable and valid in studies of psy-
chiatric patients.

The SCL emphasizes the degree of negative (dis-
tressing) symptoms only and contains several
items covering physical symptoms considered to
reflect psychological distress and psychopatholo-
gy. GHQ, on the other hand, also includes positive
questions. The respondents are requested to com-

pare their current status with their normal situation. The shorter versions of the GHQ avoid using physical symptoms as indicators of distress. Thus GHQ is probably more suitable and valid than SCL in a sample of patients admitted to the trauma unit of a surgical department. The scoring procedure of GHQ is also more easy to handle in a clinical setting than the SCL scoring procedure.

No studies have been published assessing the suitability and validity of the GHQ in a sample of trauma patients, however. The GHQ has been validated in several studies on in- and out patients seen in general hospitals and primary care settings (5, 7–10). The GHQ has also been found to be sensitive to change in psychological states following distressing events (7, 10, 11). Compared to a sample of trauma patients, the patients included in these studies are generally older and the samples include more females than seen in a trauma unit.

Accordingly, in a most recent review of the literature of GHQ by Goldberg and Williams (5), no information is provided about the sensitivity and specificity of the GHQ at different points of time in relation to the accident. This information is needed in order to use the GHQ as a tool for estimating the prevalence of psychopathology in a sample of trauma patients.

We also need to know the potential of the GHQ to identify correctly single trauma patients with psychopathology and severe distress. Used in this way, we must know which scoring procedure of the GHQ provides the optimal balance between sensitivity, specificity and positive predicting power. In order to be able to screen for any deterioration in the patient's psychosocial situation during the course of adaptation after injury, however, the clinician and researcher are interested in the performance of the GHQ in measuring change across different time periods.

The aim of the present study is to address these aspects of the 20-item version of the GHQ in a sample of hospitalized injured adults followed longitudinally for a mean period of 28 months.

Material and method

The present study of the validity of the GHQ is part of a study of psychosocial aspects of accidentally injured patients admitted to the Department of Surgery, Unit for Trauma Surgery, Akershus Central Hospital, Norway. One hundred and twelve patients aged 15–69 years were included in the study. The majority of the patients were males (76%) and below 40 years of age (mean: 36.2 years; SD 18.8 years), corresponding to a typical accident population seen in a trauma unit. The material has been bescibed in detail in another paper (4).

Design of the study

The GHQ-20 was given to the patients during the hospital stay and twice during follow-up. During the hospital stay, the patients were given a package of paper and pencil tests after a psychiatric interview. The GHQ was included in the test package. When filling in the GHQ the patients were instructed to consider only the fortnight preceding their admittance to hospital. Since cognitive deterioration due to affected physical health may influence test performance, no tests were given before the patients' ability to read, concentrate and pay attention was reasonable. This capacity was evaluated during the psychiatric interview. Since the majority of the patients had sustained only minor injuries, their hospital stay was short. Thus about half the patients were tested within the first two days following admittance to the ward.

After discharge from hospital the patients answered a mailed questionnaire including the GHQ 6–9 months following the injury. Twenty-eight months (mean) following the injury the patients again answered a mailed questionnaire including the GHQ.

Subjects

One hundred and six patients (96.4%) returned the GHQ during the hospital stay; 93 (83%) at 9 months following their accident and 110 (98.2%) at the final follow-up. Four patients lost during the hospital stay answered the GHQ at follow-up. In addition 6 relatives handed in the GHQ during the acute stage and 14 at follow-up. A few patients staying more than a fortnight in hospital filled in GHQ several times during the acute phase, giving a total of 339 questionnaires.

Eighty-nine patients returned the GHQ on all three occasions; 105 patients returned the GHQ at least twice; 91 at the acute stage and at nine months; 92 at nine months and at follow-up and 102 both during the hospital stay and at follow-up.

Psychiatric assessment

During their stay in hospital the patients were subject to a semistructured interview including the Comprehensive Psychopathological Rating Scale (12). At the time of the final follow-up, the patients were subject to another psychiatric interview including the CPRS and Barons Schedule for Interviewing Borderlines (SIB). The interview assessed both the incidence and prevalence of psychiatric diagnosis according to both ICD and the DSM-III systems of classification (2, 4).

In order to reach the most correct diagnosis, information available through medical records was carefully examined by the author. Information about previous medical and psychiatric disorders associated with sickleave from 18 years of age was obtained from the National Insurance Offices (Trygdekontorene). (Permission to receive this information was given by the National Insurance Administration – Rikstrygdeverket). This information and the findings of the clinical psychiatric examination were used to establish the psychiatric diagnosis. All the diagnoses were checked by a second psychiatrist (2, 4).

The definition of a "psychiatric case" is crucial for the prevalence of mental disorders in a population. The provision of diagnostic criteria like the DSM-III makes casefinding less biased by theoretical orientation or subjective definitions of what constitutes a psychiatric disorder. However, even the DSM-III system of classification may be difficult to use on borderline cases, since the emphasis is on categorical and not dimensional classification (13). Examples of diagnostic categories with poorly defined criteria are "psychological factors affecting physical disorders" and "adjustment disorders". In a sample of trauma patients, these disorders may be expected to occur rather frequently. The approach taken to these classification problems will thus affect the prevalence and thus the sensitivity,

specificity and positive predictive power of the GHQ. One way to meet this problem is to compute sensitivity, specificity and positive predicting power for the GHQ with and without the clear-cut cases (broad versus narrow use of DSM-III). This approach was chosen for this study.

Scoring procedure of the GHQ

The GHQ-20 contains the following items from the GHQ-60: 7, 14, 21, 26, 28, 30, 35, 36, 39, 40, 42, 43, 46, 47, 49, 50, 51, 54, 55, 58. The possible scoring range depends on the scoring procedure.

The answers to each item may be treated as a multiple-response scale or "Likert scale" and have weights assigned to each position (0-1-2-3). This gives a possible range for the total GHQ-20 score of 0-60. When measuring change with the GHQ, this Likert score procedure has often been applied.

A much easier way to score the responses to the items is to treat the GHQ as a bimodal response scale. According to this procedure only pathological deviations from normal will signal possession of the item. This method of scoring has been called "GHQ-scoring". The possible range for scores is 0-20. This scoring procedure is easy to compute in clinical practice but the advantage is gained only at the cost of losing information. However, when using the GHQ as a screening test, the GHQ scoring has proved to provide acceptable values for sensitivity and specificity. Usually a cutting point between 3 and 4 has been used (5). In this paper a score of 4 or more is called GHQ-case score unless otherwise specified.

Goodchild & Duncan-Jones (14) have proposed a new scoring procedure. They argue that the response "no more than usual" to an item describing pathology (example: "lost much sleep over worry" or "been feeling nervous and strung-up all the time") should be treated as an indicator of chronic illness. In the 20-item version of the GHQ, ten items have this response as a possibility. This scoring procedure does not change the possible range for scores compared to the GHQ scoring procedure. A recent study has

Table 1a

The sensitivity (sens), specificity (spec) and positive predicting power (PPP) of the 20-item version of the General Health Questionnaire with regard to psychiatric disorder in a prospective study of 113 accidentally injured adults.

Filled in during the hospital stay											
GHQ-score Scoring				Likert score Scoring				Chronic score Scoring			
Cut-off	Sens	Spec	PPP	Cut-off	Sens	Spec	PPP	Cut-off	Sens	Spec	PPP
(2/3)	62%	83%	62%	(17/18)	75%	65%	44%	(9/10)	58%	77%	47%
(3/4)	48%	91%	67%	(22/23)	45%	91%	65%	(11/12)	38%	93%	69%
(4/5)	45%	95%	76%	(27/28)	28%	99%	89%				

Table 1b

The sensitivity (sens), specificity (spec) and positive predicting power (PPP) of the 20-item version of the General Health Questionnaire with regard to psychiatric disorder in a prospective study of 113 accidentally injured adults.

Filled in at follow-up (mean 28 months post accident)											
GHQ-score				Likert score				Chronic score			
Cut-off	Sens	Spec	PPP	Cut-off	Sens	Spec	PPP	Cut-off	Sens	Spec	PPP
(3/4)	73%	79%	56%	(23/24)	69%	81%	59%	(10/11)	76%	77%	63%
(4/5)	65%	84%	64%	(27/28)	56%	92%	75%	(11/12)	63%	82%	

suggested that this scoring procedure may identify a higher proportion of the middle-aged (15).

The GHQ may be used in three different ways: (1) The percentage of a population with a high score on the GHQ may be used to *estimate the prevalence* of psychiatric disorders in the population. For this purpose we need to know the sensitivity and the specificity of the GHQ in a similar population.

(2) A persons' total score on the GHQ may be used to *identify persons* with psychiatric disorders. The decisive factor for this clinical use is the GHQ's ability to identify true cases according to widely used classification systems of mental disorders like the Diagnostic and Statistical Manual of Mental Disorders, 3rd edition or DSM-III-Revised and the International Classification of Diseases, section for mental disorders (ICD-9). This ability is reflected in the tests' positive predictive power. It is also of clinical interest to know if negative test results really reflect true negative cases. This ability is reflected in the negative predictive power (16).

(3) The GHQ may be given on several occasions to the same person in order to *measure change* in

mental state following upsetting or distressing events like accidents and injuries. Used in this way the change in total score is the crucial dimension. If valid for this use the difference between the score before the event and after the event should identify those with and those without change in mental state.

In this study the distribution curves of the scores for all three scoring procedures were made in order to identify the most valid scoring system and the most appropriate cutting points.

Reliability

(1) Split-half reliability was computed by correlating the sum score of the first ten items with the sum score for the last ten items. No items were paired. In the acute phase the coefficient was 0.81; in the first follow-up the coefficient was 0.86 and in the final follow-up 0.81. For all questionnaires filled in (N = 339) the split-half coefficient was 0.84.

(2) The internal consistency (Chronbach's alfa) for the GHQ-20 during the three different occasions was 0.81, 0.86 and 0.91 respectively. These findings confirm the good psychometric properties of the GHQ.

Statistics

The Wilcoxons signed rank test for matched pairs was used to assess significance of the differences between the scores on two different occasions. Correlations were calculated by means of Pearsons r.

Results

Identification of psychiatric cases

The sensitivity, specificity and positive predicting power of narrow DSM-III diagnosis before the accident and at follow-up were calculated for all three scoring-systems.

The sensitivity and specificity are better in the follow-up than in the acute phase (Table 1). The optimal sensitivity, specificity and positive predictive power of the simple GHQ-scoring procedure did not differ substantially from the two other scoring systems.

The most appropriate cutting point depends on the aim of the screening. The distribution curves suggested a cutting point between 3 and 4 using the GHQ-system, 10 and 11 using the Chronic score, and 23 and 24 using Likert scoring is reasonably appropriate at follow-up. Accordingly, these cutoff points were used in the succeeding analysis.

The sensitivity and specificity of a screening test depend basically on the definition of what constitutes a psychiatric case. In Table 2 the sensitivity and specificity of the GHQ-20 are shown classifying patients according to ICD-9 and a narrow and more broad use of DSM-III. The latter concept includes psychological factors affecting physical disorders ("psychosomatic disorders"). The sensitivity and specificity are slightly better for the ICD-9 diagnosis. The more complicated Likert and Chronic scoring procedures, however, do not increase the sensitivity and the specificity of the GHQ very much compared to the results using the simple conventional scoring procedure.

Chronic-score seems best at classifying DSM-III disorders in a narrow sense and ICD-9 disorders are best predicted by Likert scoring.

Eight patients were falsely classified as DSM-III cases according to the GHQ-scoring procedure during the hospital stay. Two of these patients did qualify for a ICD-9 diagnosis, however, and one patient was in a state of bereavement (Table 3).

The patients missed by the GHQ (false negative) at follow-up are listed in Table 4. The majority of these cases are patients with personality disorders usually not associated with strong feelings of subjective affects.

Measuring change

The validity of the GHQ in detecting change in mental state is of practical interest for clinicians.

Table 2

A prospective study of accidentally injured adults. The sensitivity, specificity and positive predicting power of the 20-item version of General Health Questionnaire in relation to definition of a case.

	DSM-III-N		DSM-III-B		ICD-9	
	Sens	Spec	Sens	Spec	Sens	Spec
GHQ-scoring	73%	79%	73%	79%	75%	84%
Likert scoring	69%	81%	73%	84%	75%	88%
Chronic scoring	76%	77%	76%	77%	80%	89%

	DSM-III-N PPP	DSM-III-B PPP	ICD-9 PPP
GHQ-scoring	56%	63%	71%
Likert scoring	59%	65%	73%
Chronic scoring	72%	58%	65%

DSM-III-N = diagnoses without clear-cut inclusion criteria are avoided (narrow diagnosis)
DSM-III-B = diagnoses lacking clear-cut inclusion criteria are included. (Broad diagnosis). See text.
ICD-9 = diagnoses according to the Norwegian use of the system.

Table 3

A prospective study of accidentally injured adults. False positive psychiatric cases according to GHQ-20 screening using DSM-III as a criterion for case-ness. The GHQ-20 was filled in during the hospital stay.

GHQ Score	ICD diagnosis/psychological characteristic
15	Character neurosis
6	Uncomplicated bereavment
5	Character neurosis
5	Histrionic personality traits
4	Histrionic and asthenic personality traits
4	Normal teenager girl, current conflict about dependence and independence
4	Normal teenager girl, current conflict about dependence and independence

Figure 1 shows that the distribution of GHQ-scores has moved towards higher values at the final follow-up compared with the pre-accident scores. The differences between the GHQ score the fortnight preceding admittance to hospital ("acute") and the first follow-up, and the acute score and second follow-up score were statistically significant ($P < 0.001$). There was no statistically significant difference between the scores at the first and second follow-up (P = 0.7). These findings are reflected in all three scoring systems.

Further analyses showed that the patients who had permanently higher case-scores at follow-up than pre-accident were among those most severely affected by the injuries from a psychosocial point of view (Table 5).

The subjects who had a pathological score after answering the GHQ on all three occasions

Table 4

A prospective study of accidentally injured adults. False negative cases according to GHQ-20.

The DSM-III diagnosis of patients not identified by a case score on the GHQ-20 at follow-up (N = 7). The number of diagnoses are above seven since some patients qualify for more than one diagnosis.

Axis 1	
Substance abuse	4
Organic mental disorder	1
Anxiety disorder (chronic)	1

Axis 2	
Antisocial personality disorder	3
Avoidant personality disorder	2
Histrionic personality disorder	1
Dependent personality disorder	1
Mixed personality disorder	1

(N = 9) qualified for a major DSM-III diagnosis at the final follow-up. Of those who scored as a case on two occasions (N = 19), 63% had a DSM-III diagnosis at follow-up.

Of the 28 patients scoring as a case at least once, 18% had a DSM-III diagnosis at follow-up. Only 10% of the persons with no case score at the three different times (N = 33) qualified for a DSM-III diagnosis on the final follow-up.

There was no correlation between GHQ-score and time lag between filling in the GHQ and the psychiatric examination on the different occasions (range of r = +0.04 to -0.12) whatever the scoring system used.
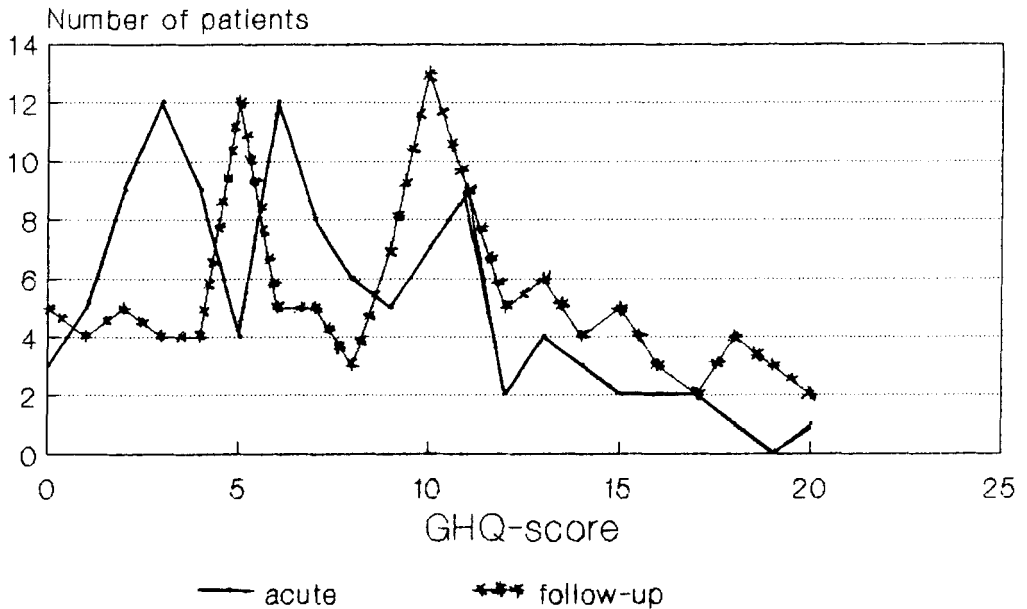
Discussion

Estimating prevalence

The study indicates that GHQ-20 is a valid screening instrument for clinical use among accidentally injured patients. However, during their hospital stay the sensitivity of the GHQ is less than optimal. A sensitivity of 62% and a specificity of 83% may be used to assess the prevalence of DSM-III diagnoses the fortnight preceding the injuries if GHQ scoring is applied with a cutting point between 2 and 3. This cutting point is below the usual cutting point of 3 and 4; the reason for this is probably the method applied. Patients with mental problems did in some cases underestimate their level of distress experience before the accident. In the light of the present trauma experience, the immediate past in some cases appeared more pleasant, thereby reducing the sensitivity of the

Fig. 1. GHQ-20 scores - acute stage and at follow-up.

test. This observation needs to be replicated before the values for sensitivity and specificity can be used in other populations.

In a follow-up situation, the study suggests a sensitivity and specificity of 75% and 84% using GHQ scoring procedure and a cutting point between 3 and 4. This is similar to the sensitivity and specificity found in studies on patients seen in general medical practice (5) and not better than screening questionnaires designed specially for post-traumatic problems of any kind (16). Usually these patients are older than the present population. One may wonder why the prevalence, and thus the sensitivity and specificity, among accidentally injured patients reach similar levels despite their younger age. The reason for the present incidence is probably the increased psychopathology in the present population (2, 4).

Considering the cost and benefits of a complicated versus a simple scoring procedure, the present finding does not support the value of using a more complicated scoring procedure in clinical

settings as suggested by Goodchild & Duncan-Jones (14). This finding may, however, be related to the younger age group in question in the present study (15).

This study suggests that the sensitivity and specificity in the present population depend on four main factors which may vary from study to study: The definition of mental disorder; the length from accident or injury until the GHQ is filled in; the scoring system used and the cutting point used. Similar findings have been reported by others studying different type of patients (17–19).

Identifying patients with mental problems

Used as a screening instrument, this study shows that GHQ-20 may identify 73–80% of the patients with psychiatric disorders. However, the positive predictive power determines its usefulness as an instrument for identifying individual patients.

Table 5

A prospective study of adults (N = 89) admitted to a trauma unit due to accidental injuries.
Distribution of patients with normal GHQ-20 score (0–3) or pathological score (4–20) on three different occasions: Acute during the hospital stay; first follow-up = 6–9 months following admittance to hospital; final follow-up = 28 months following admittance to hospital.

	Acute*		6–9 months		28 months		Total
	Non	Case	Non	Case	Non	Case	
No change (n = 42):	33		33		33		
		9		9		9	42
Unstable (n = 35):	13		13			13	
	11			11	11		
		3	3			3	35
		2	2			2	
		6		6	6		
Permanent change (n = 12):	9			9		9	
		3	3		3		12
	66	23	54	35	53	36	
Total	89		89		89		89

*The GHQ was filled in during the hospital stay but the patients were asked to consider the last two weeks preceding the accident.

Using a conventional cut-off point, the PPP is only modest. By increasing the cut-off point, the PPP may be raised but only at the cost of a low sensitivity. This finding corresponds to other studies of hospital patients and is logical in view of the fact that the single use of GHQ was constructed as a screening instrument having a two-phase design (5).

The present study does, however, indicate that repetitive use of the GHQ may in fact be a very good measure of more chronic and stable psychiatric problems, including the more severe axis II disorders, except for antisocial personality disorder. Future research is needed to replicate this finding.

The unidentified cases are in fact the same cases found to be false negatives in other studies. When applying questionnaires like the GHQ it has been observed that more false negatives occur among patients with personality disorders and substance abuse (8) and as well as among some patients with chronic disorders currently in a good phase (20). Similarly it is well known that the false positives are most often persons with histrionic traits who self-dramatize or persons under psychological stress.

This was also seen in the present study, although some of these persons would be con-sidered to have psychiatric problems according to Norwegian psychiatric practice (Table 5).

A focus on psychiatric cases only may, however, not be all that relevant for clinical practice. It has been observed that medical and surgical patients experiencing psychological distress during their hospital stay may run the risk of suffering from more long-term complications (21). The socalled false positives according to psychiatric caseness, were, however, identified by the GHQ in addition to persons with psychopathology in the present study. Thus a score of 4 or more on the GHQ may give useful clinical information to the treating surgeon about patients at risk. From this point of view, the term false positive according to DSM-III or ICD-9 is less relevant for the surgeon, whereas a score above 3 on GHQ-20 should be very useful information for the surgeon.

However, if the aim is to detect persons with psychopathology only, the most appropriate scoring procedure and cutting point depend on the aim of the screening. If the aim is to carry out a broad screening programme highly sensitive tests can be particularly useful since negative test results may be used to exclude persons for further screening (22). This is reflected in NP, which may be expected to be between 80 and

90% in a population similar to the present one. Considering the complicated procedure necessary for calculating the Likert or Chronic score, the present study suggest that the simple GHQ scoring with the usual cutting point between 3 and 4 is satisfactory. Notice that the sensitivity of the GHQ-20 reaches 100% if a case score is seen on three subsequent occasions over a two-three year period.

Tests highly specific but more limited in sensitivity are more appropriate for confirmatory use after initial screening. This is reflected in the PP. This study may indicate that used in this way Likert scoring may be most appropriate using a rather high cutting point like 27/28.

Measuring change

The present findings indicate that GHQ should be a valid and simple test for measuring change in psychological balance. This has been found in other studies too using the GHQ and the Schedule for Affective Disorders and Schizophrenia, life time version (23). However, although the more *severe* changes to the worse in a group of patients may be detected by simple GHQ scoring procedure, Likert scoring procedure may be slightly better to detect minor change within the normal range of scores.

In this study special attention was paid to the validity of psychiatric diagnosis and the assessment of psychosocial distress. However, a six month period between the patient's reply to GHQ and the final follow-up interview occurred in some cases. One may argue that this may reduce the sensitivity of the clinical interview at least for short term mental distress and minor disorders of short duration like adjustment disorders. This may falsely reduce the validity of the GHQ (18). A life event questionnaire was mailed together with the tests at the first follow-up (4). When conducting the interview at final follow-up, special attention was paid to the occurrence of life events during the observation period by consulting the life event questionnaire at interview. Also information from the Health Insurance System helped in assessing the psychiatric state of the patients during the observation period. Thus probably

the timelag between the test and the interview in some patients has not had decisive influence on the results.

Conclusion

The 20-item version of the General Health Questionnaire is a valid screening instrument for psychopathology and distress occurring among accidentally injured persons. Repetitive use of the GHQ-20 may be used to identify permanent change to the worse in persons suffering from injuries. Considering the importance of taking emotional factors into account in rehabilitation and how easy it is to obtain a GHQ score, the routine use of GHQ-20 in the follow-up of accidentally injured persons should be encouraged.

References

1. Malt U. Biopsychosocial aspects of accidents. In, Murison R & Ursin H, eds, Biological and psychological basis of psychosomatic disease (Advances in the Biosciences volume 42) Oxford: Pergamon Press 1983, 103–113.
2. Malt U, Myhrer T, Blikra G, Høivik B. Psychopathology and accidental injuries. Acta Psychiatr Scand 1987:76:261–271.
3. Malt U, Blikra G, Høivik B. The 3-year biopsychosocial outcome of 551 hospitalized accidentally injured adults. Acta Psychiatr Scand 1989:80:Suppl. 355, 84–93.
4. Malt U. The long-term psychiatric consequences of accidental injury. Br J Psychiatry 1988:153810–818.
5. Goldberg D, Williams P. A user's guide to the General Health Questionnaire. Windsor: NFER-Nelson 1988.
6. Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH. Covi L. The Hopkins Symptom Checklist (HSCL). Mod probl Pharmacopsychiat 1974:7:79–110.
7. Merskey H, Lau CL, Russel ES, Brooke RI, James M, Lappano S et al. Screening for psychiatric morbidity. The pattern of psychological illness and premorbid characteristics in four chronic pain populations. Pain 1987:30:141–157.
8. Vazques-barquero JL, Acero JAP, Martin CP & Ochoteco A. The psychiatric correlates of coronary pathology: Validity of the GHQ-60 as a screening instrument. Psycholog Med 1985:15:589–596.
9. Lobo A, Perez-Echeverria M-J, Jimenez-Aznarez A, Sancho MA. Emotional disturbance in endocrine patients. Br J Psychiatry 1988:152:807–812.
10. Sharp DJ. Validation of the 30-item General Health Questionnaire in early pregnancy. Psycholog Med 1988:18:503–507.
11. Vachon MLS, Lyall WAL, Rogers J, Freedman letofsky & Freeman SJJ. A controlled study of self-help intervention for widows. Am J Psychiatry 1980:137:1380–1384.

12. Åsberg M, Montgomery SA, Perris C, Schalling D & Sedvall GA. Comprehensive Psychopathological Rating Scale. Acta psychiat Scand, 1978: suppl. 272.

13. Malt U. Five years experience with the DSM-III system in clinical work and research: some concluding remarks. Acta Psychiatr Scand 1986:73: supplement. 328, 76–84.

14. Goodchild ME & Duncon-Jones P. Chronicity and the General Health Questionnaire. Brit J Psychiatry 1985:146:55–61.

15. Huppert FA, Gore M, Elliott BJ. The value of an improved scoring system (CGHQ) for the General Health Questionnaire in a representative community sample. Psycholog Med 1988:18:1001–1006.

16. Malt U, Blikra G, Høivik B. The Late Effect of Accidental Injury Questionnaire. Acta Psychiatr Scand 1989:80:Suppl. 355, 113–130.

17. Dean. C, Surtees OG & Sashidharan SP. Comparison of research diagnostic systems in an Edinburgh community sample. Brit J Psychiatry 1983:142:247–256.

18. Duncan-Jones P & Henderson S. The use of a two-phase design in a prevalence survey. Social Psychiatry 1978, 13, 213–237.

19. Goldberg D. Identifying psychiatric illness among general medical patients. Brit. Med J 1985:291:161–162.

20. Finlay-Jones RA, Murphy E. Severity of psychiatric disorder and the 30-item General Health Questionnaire. Brit J Psychiatry 1979:134:609–616.

21. Cromwell RL, Butterfield EC, Brayfield FMM & Curry JJ. Myocardial infarction: Reaction and recovery. St. Louis: C V Mosby 1977.

22. Baldessarini RJ, Finklestein S & Arana GW. The predictive power of diagnostic tests and the effect of prevalence of illness. Arch Gen Psychiatry 1983:40:569–573.

23. Kessler LG, Cleary PD & Burge JD. Psychiatric disorders in primary care. Arch Gen Psychiatry 1985:42:583–587.