The FIP Study: a randomised, controlled trial of screening and recognition of psychiatric disorders

Kaj Sparle Christensen, Tomas Toft, Lisbeth Frostholm, Eva Ørnbøl, Per Fink and Frede Olesen

SUMMARY

Background: *Research on questionnaires as screening tools for psychiatric disorders has yielded conflicting results.*

Aim: *To examine the effect of a routinely administered questionnaire on recognition of common psychiatric disorders in general practice.*

Design of study: *Randomised controlled trial.*

Setting: *Twenty-eight general practices in Aarhus County, Denmark.*

Method: *Thirty-eight general practitioners (GPs) and 1785 consecutive patients, aged 18–65 years old, presenting with a new health problem, participated. Before consultation, patients were screened using a brief screening questionnaire (SQ) including somatisation, anxiety, depression, and alcohol abuse scales. Patients were randomised to one of two groups: 900 questionnaires were disclosed and scored by the GPs, 885 were blinded. A stratified subsample of 701 patients was interviewed after the consultation using a standardised psychiatric research interview (SCAN).*

Results: *Overall the GPs' recognition rates were 14% (95% confidence interval [CI] = -2 to 30) better for depression and 35% (95% CI = 2 to 68) better for alcohol problems when SQs were disclosed. Recognition rates for anxiety improved 8% (95% CI = -9 to 26) overall. In the case of somatoform disorders, disclosure showed no effect overall. Among those with high SQ scores, however, disclosure increased recognition rates on any mental disorder evaluated.*

Conclusion: *This study demonstrated limited usefulness for routine screening for common psychiatric disorders. However, findings suggest that the SQ may be useful for case-finding among a subgroup of patients with high SQ scores.*

Keywords: *mental disorders; randomised controlled trials; psychiatric status rating scales; screening tests; recognition (psychology).*

K Sparle Christensen, MD, research assistant, The Research Unit for Functional Disorders, Aarhus University Hospital and The Research Unit for General Practice, Aarhus University. T Toft, MD, research assistant; L Frostholm, MSc, research assistant, psychologist; E Ørnbøl, MSc, statistician; P Fink, PhD, DMSc, associate professor, director, The Research Unit for Functional Disorders, Aarhus University Hospital. F Olesen DMSc, FRCGP, professor, head of department, The Research Unit for General Practice, Aarhus University, Denmark.

Address for correspondence

Kaj Sparle Christensen, Barthsgade 5.1, DK-8200 Aarhus N, Denmark. E-mail: kspar@akh.aaa.dk

Submitted: 30 August 2002; Editor's response: 16 January 2003; final acceptance: 28 July 2003.

Introduction

PSYCHIATRIC disorders, such as somatoform disorders, anxiety, depression, and alcohol abuse, are prevalent in general practice, but often go unrecognised.[1,2] Attempts to improve recognition by using psychiatric screening questionnaires have yielded conflicting results. A systematic review of randomised controlled trials on routine screening for depression and anxiety in non-psychiatric settings demonstrated no increase in overall recognition; although the detection rates for depression improved among patients with high scores.[3] Diagnostic instruments; for example, DSM-IV-PC, PRIME-MD and PHQ, have been introduced in primary care. No controlled trials on diagnostic instruments have previously been published, but results from uncontrolled trials indicate that these instruments can be useful.[4-6]

The aim of the study was to examine if routine screening for somatisation, anxiety, depression, and alcohol abuse would improve the general practitioners' (GPs') recognition rates compared to the results of a standardised psychiatric interview.

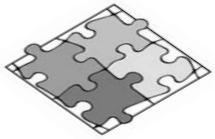
Method

Participants

Aarhus County has a population of 600 000 people living in rural or metropolitan areas. The county is served by 431 GPs working in 271 practices. All GPs were invited to participate in a randomised controlled trial on recognition and treatment of functional illness in primary care (the FIP Study). Thirty-eight (8.8%) GPs working in 28 practices accepted the invitation to participate. Half of the participating GPs were randomised to an educational programme on somatisation.[7] This study on psychiatric screening was part of the FIP Study and includes all of the participating GPs.

Included in the study were consecutive patients aged 18–65 years presenting with a new health problem during a 3-week period (3 March–1 May 2000). Patients of non-Scandinavian descent, patients who could not speak or read Danish, and patients who were too ill or demented to read and fill in the questionnaires, were excluded. Only patients enrolled in the National Health Care Programme, which covers 98% of the Danish population, were included. All participating patients were registered with one GP, and consulted this doctor and/or practice primarily. Nearly all specialised treatment, including hospital admission, requires GP referral, except in emergency cases. The Danish healthcare system is almost entirely tax financed and most medical care is free.

HOW THIS FITS IN



What do we know?

Much psychiatric morbidity goes unrecognised in primary care. Results from randomised controlled trials on psychiatric screening in primary care are conflicting.

What does this paper add?

This study demonstrated limited usefulness from *routine* screening for common psychiatric disorders. However, findings suggest that the screening questionnaire may be useful for case-finding among a subgroup of patients with high questionnaire scores.

Interventions

Screening questionnaire. All of the patients included in the study were screened in the waiting room before consultation. A one-page screening questionnaire (SQ) was used, consisting of various rating scales: for somatisation this was the SCL-90R somatisation subscale (SCL-SOM)[8] and Whiteley-7,[9] for anxiety and depression this was the SCL-8,[10,11] and for alcohol abuse this was the CAGE.[12] After completion, patients were randomised to either blinding or disclosure of the SQ information to the GP. All of the GPs were instructed on how to rate the questionnaire using a handy scoring template at the beginning of the consultation. Scores were expressed as positive predictive values (PPVs) on somatisation (symptom and illness worry score), mental disorder in general, depression, and alcohol dependence/abuse. The completion of the SQ usually took 2–5 minutes and the rating less than 1 minute. Immediately after consultation, the GPs completed a questionnaire on their own assessment. All GPs were free to use a handout stating ICD-10 criteria for somatoform disorder, anxiety, depression, and alcohol abuse.

Outcomes

The main outcome measures were the GPs' recognition of somatoform disorders, anxiety, depression, and alcohol abuse compared to the results of a psychiatric research interview. The GPs were asked to classify the main problem presented in one of the five following categories: physical disease, probable physical disease, medically unexplained symptoms, mental illness, and no physical problem. Later, this categorisation was dichotomised into 'physical disease' (first two categories) or 'somatisation'. The GPs were specifically asked whether the patient suffered from a significant psychiatric disorder, i.e. anxiety, depression or alcohol abuse.

After the consultation, a stratified sample, including every ninth eligible patient and all patients with high SQ scores, was selected for a semistructured standardised psychiatric interview (SCAN).[13] High scoring patients were defined as dichotomised scores on SCL-SOM >3, or Whiteley-7 >1, or SCL-8 >1, or CAGE >1. Among 894 selected patients, 193 (21.6%) declined, leaving 701 for interview. The SCAN generates ICD-10 diagnoses[14] and has been used as a gold standard. Six psychiatrically trained physicians, certified at the Copenhagen WHO-SCAN Training Centre, conducted the interviews. Interviewers were blinded to screening results and randomisation status.

Sample size

Sample size calculations targeted the educational programme. A simulation study was performed in order to assess how many GPs and patients would be required. A significant difference ($P<0.05$) from a chosen scenario in health between intervention and control patients was ascertained with a power of 68% upon inclusion of 300 patients with a positive somatoform diagnosis (15 GPs per group, 10 patients per GP), or with a power of 80% upon inclusion of 400 patients with a positive somatoform diagnosis (20 GPs per group, 10 patients per GP). Prevalence results from other studies suggest that 2000–2500 patients were needed to reach a sample size of 400 patients with a positive somatoform diagnosis.[1]

Randomisation and blinding

The medical secretaries were instructed on registration and inclusion procedures. Having obtained the patients' informed consent, the secretary broke a concealed non-transparent envelope with SQ and GP questionnaires. Having completed the SQ, patients were randomised to have the SQ disclosed or blinded to their GP. A colour code on the GP questionnaire clearly stated whether the secretary should hand over the SQ to the GP (disclosure) or whether to return it to the envelope (blinding).

Data processing

Data from questionnaires and SCAN interviews were gathered using TELEform formulas. Diagnoses from SCAN interviews were processed by SCAN I-Shell (computer assisted personal interviewing application for SCAN 2.1) except for somatoform diagnoses, which were reprocessed separately with computer algorithms strictly according to ICD-10 criteria.

Statistical analysis

All randomised patients were included. Unanswered questions on the SQ were automatically scored zero, based on the assumption that patients only responded to questions relevant to them. Our analysis was based on random allocation using complete datasets only. Statistical analysis was performed using STATA 7.0 for Windows and SPSS 10.0 for Windows. We investigated differences in baseline characteristics using unpaired *t*-tests and χ^2 tests. Sensitivity (SE), specificity (SP), correctly classified (CC), and relative odds ratio (ROR) were analysed using weighted logistic regression, thus correcting for skewness introduced by the stratified sampling procedure.[15] Absolute benefit increase (ABI), relative benefit increase (RBI) and number needed to test (NNT) were analysed using general linear modelling (GLM, family: Bernoulli, link: Identity) with the same weights to correct for the two-phase sampling procedure. Subgroup analyses on GPs and included patients were performed using weighted logistic regression. No explanatory variables were included in the logistic models for SE, SP and CC. For the analysis of ROR the explanatory variables were SQ, GP diagnosis and their interaction. In the GLM models SQ was the only explanatory variable.

Ethics

This study was approved by The Science Ethics Committee in the Aarhus County, the Data Surveillance Authority and the Scientific Research Evaluation Committee of the Danish College of General Practitioners.

Results

Participant flow

Among the 2424 patients assessed for eligibility, 227 met exclusion criteria, 274 patients declined, and 138 could not participate for other reasons; for example, time pressure in clinic and patients not carrying reading glasses. Each patient was included only once. A total of 1785 (81.2%) patients joined the study (Figure 1).

Baseline data

Participating GPs had practiced family medicine for fewer years than non-participating GPs (mean = 10.3 versus 14.1 years, likelihood ratio test [LR-test] < 0.005), and more often reported having participated in longer courses (more than 3 days) in communication skills or psychological therapy (52.8% versus 39.5%, LR-test < 0.05). Length of postgraduate psychiatric training, types of practice, and the GPs' age and sex were not found to have any statistically significant influence on recognition rates.

Declining patients had a mean age of 42.2 years compared to 38.8 years among included patients ($P<0.001$, *t*-test), whereas we found no statistically significant sex differences. No significant differences were found in age or sex between included patients and patients not participating for other reasons. Differences in randomisation status (details in Table 1) were small; disclosure patients were slightly older ($P = 0.09$, *t*-test).

Among the 894 patients selected for the SCAN interview, 701 (78.4%) accepted. Patients with low scores, younger patients, and men refused the interview more often than other groups, but the differences were small. Inter-rater agreement on psychiatric diagnosis was high ($\kappa = 0.86$) (Toft *et al*, unpublished data, 2002).

Outcomes and estimation

Data weighting was based on a stratified sample of 615 patients with high SQ scores and 84 with low scores (Table 2). Weighted estimates (Tables 3 and 4) suggested increased recognition of alcohol abuse, depression, and anxiety by disclosure, but levels only reached statistical significance for the first mentioned. Disclosure obviously improved overall correctness for any diagnosis of anxiety, depression, and alcohol abuse (ROR = 2.8, 95% CI = 1.1 to 7.3). Surprisingly, GP sensitivity to somatisation declined and specificity rose when SQs were disclosed.

Disclosure of SQs on high scores significantly improved recognition rates of all mental disorders evaluated (Table 5).

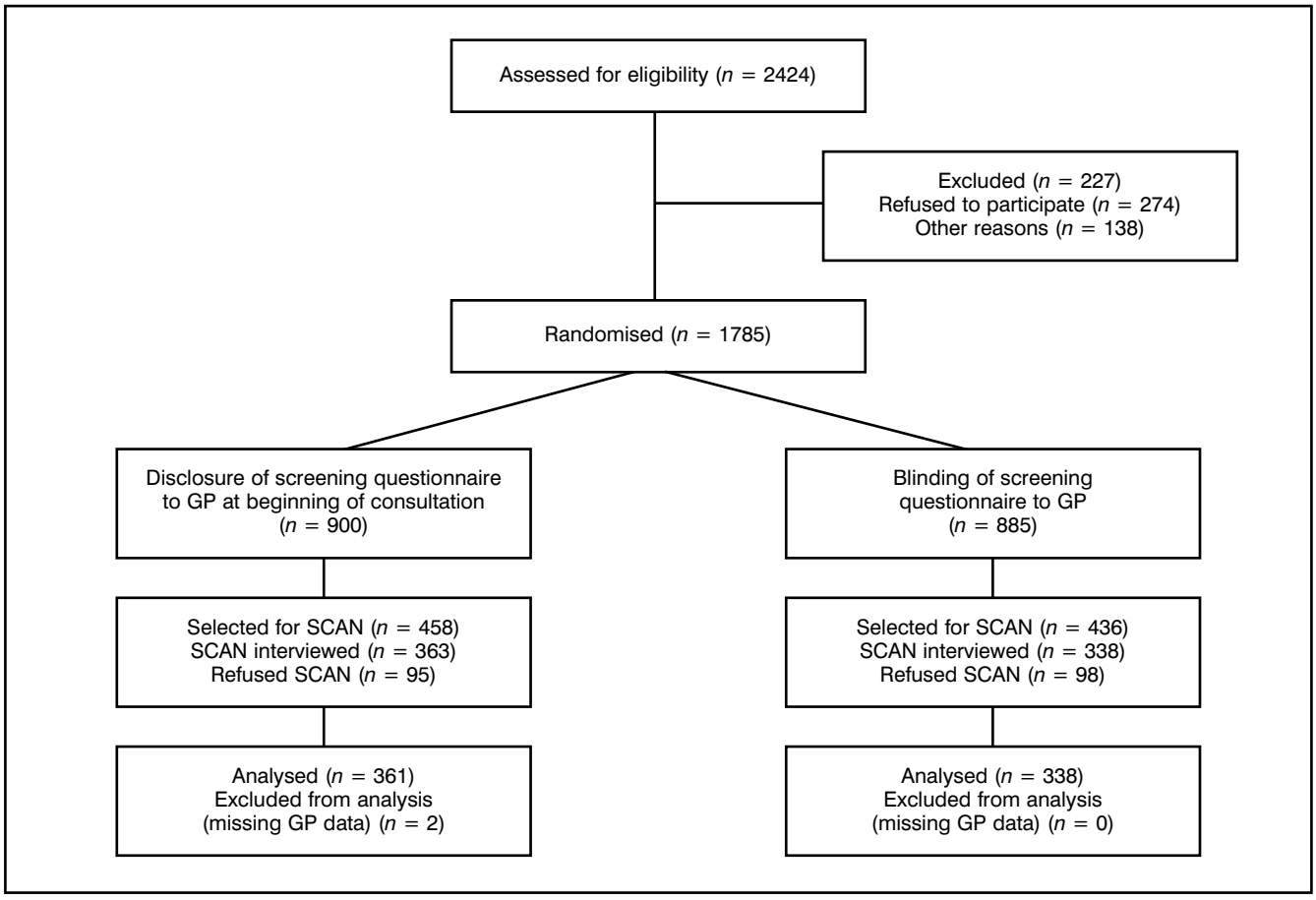


Figure 1. Patient trial flow.

Table 1. Data on randomised patients.

Variables	Blinding of SQ[a]	Disclosure of SQ	*P*-value
Number of patients	885	900	0.72
Percentage female	61	59	0.72
Mean age in years (SD[b])	38.2 (12.9)	39.3 (12.9)	0.09
High score on rating scales	377	398	0.49
Mean SCL-SOM scores (SD)	2.3 (2.3)	2.2 (2.3)	0.30
Mean Whiteley-7 scores (SD)	0.9 (1.5)	0.9 (1.5)	0.66
Mean SCL-8d scores (SD)	1.2 (2.1)	1.1 (2.0)	0.73
Mean SCL-anxiety scores (SD)	0.6 (1.0)	0.6 (1.0)	0.81
Mean SCL-depression scores (SD)	0.6 (1.3)	0.6 (1.3)	0.88
Mean CAGE-alcohol scores (SD)	0.2 (0.6)	0.2 (0.6)	0.82
GP not allocated/allocated to educational programme	439/461	433/452	0.95

[a]SQ = screening questionnaire; [b]SD = standard deviation.

Table 2. Data on SCAN interviewed patients. GPs' recognition of psychiatric disorders according to randomisation status; blinding (SQ-) or disclosure of screening information (SQ+).

ICD-10 diagnosis	Intervention	GP sensitivity (%)	GP specificity (%)	Correctly classified (%)
High scores on screening questionnaires/(*n* = 617[a])				
Somatoform disorders[b]	SQ-	49/141 (34.8)	106/160 (66.3)	155/301 (51.5)
	SQ+	70/150 (46.7)	119/164 (72.6)	189/314 (60.2)
Anxiety disorder[c]	SQ-	18/71 (25.4)	195/230 (84.8)	213/301 (70.8)
	SQ+	31/75 (41.3)	201/239 (84.1)	232/314 (73.9)
Depression, including dysthymia[d]	SQ-	37/91 (40.7)	190/210 (90.5)	227/301 (75.4)
	SQ+	50/90 (55.6)	199/224 (88.8)	244/314 (79.3)
Depression, excluding dysthymia[e]	SQ-	32/70 (45.7)	206/231 (89.2)	238/301 (79.1)
	SQ+	46/77 (59.7)	208/237 (87.8)	254/314 (80.9)
Alcohol abuse[f]	SQ-	4/16 (25.0)	278/285 (97.5)	282/301 (93.7)
	SQ+	9/15 (60.0)	283/299 (94.7)	292/314 (93.0)
Low scores on screening questionnaires/(*n* = 84)				
Somatoform disorders[b]	SQ-	3/7 (42.9)	27/30 (90.0)	30/37 (81.1)
	SQ+	3/17 (17.6)	29/30 (96.7)	32/47 (68.1)
Anxiety disorder[c]	SQ-	0/2 (0.0)	35/35 (100.0)	35/37 (94.6)
	SQ+	1/7 (14.3)	39/40 (97.5)	40/47 (85.1)
Depression, including dysthymia[d]	SQ-	1/0 (-)	36/37 (97.3)	36/37 (97.3)
	SQ+	0/1 (0.0)	45/46 (97.8)	45/47 (95.7)
Depression, excluding dysthymia[e]	SQ-	1/0 (-)	36/37 (97.3)	36/37 (97.3)
	SQ+	1/0 (-)	46/47 (97.9)	46/47 (97.9)
Alcohol abuse[f]	SQ-	0/0 (-)	37/37 (100.0)	37/37 (100.0)
	SQ+	0/0 (-)	47/47 (100.0)	47/47 (100.0)

[a]Missing GP data on 2 (0.3%) interviewed patients; [b]Somatoform disorders (F44, F45, F48.0); [c]Anxiety, exclusive specific phobias (F40.2) and somatoform disorders (F44, F45, F48.0); [d]Mood disorders, exclusive F30, F31.0–31.2, F31.7–31.9, F33.4; [e]Mood disorders, exclusive F30, F31.0–31.2, F31.7–31.9, F33.4 and 'dysthymia' (F34, F38); [f]Mental and behavioural disorders owing to use of alcohol (F10.01–F10.6).

Weighting data on high scores did not influence results decisively.

Subgroup analyses showed that GPs younger than 50 years old failed to recognise somatisation more often than GPs older than 50 years old when screening negative results were disclosed ($P = 0.01$), yet they more often correctly diagnosed any alcohol abuse disorder ($P = 0.001$). Screening disclosure did not produce statistically significant differences in recognition, which could be related to differences in GP sex, years of practicing family medicine, previous experience with communication and psychotherapy, or randomisation to the educational programme. Nor did subgroup analyses show any statistically significant differences in recognition rates as a result of screening disclosure, which could be related to patient age or sex.

Discussion

Summary of main findings

This study demonstrated limited usefulness from routine screening for common psychiatric disorders in daily clinical practice. However, findings suggest that the SQ may be useful for case-finding among a subgroup of patients with high SQ scores.

The strengths and the limitations of this study

The data suggest unbiased randomisation and successful blinding, but the accuracy of the inclusion procedure depended on the medical secretaries. Our analyses were based on random allocation and not 'intention to diagnose/treat', as 138 patients were registered as 'not participating for other reasons'. The found benefit from screening

Table 3. General practitioners' recognition of psychiatric disorders according to randomisation status; blinding (SQ-) or disclosure of screening information (SQ+). n *=701[a], weighted data = 1785.*

ICD-10 diagnosis	Intervention	GP sensitivity (95% CI)	GP specificity (95% CI)	Correctly classified (95% CI)
Somatoform disorders[b]	SQ-	36.8 (25.1 to 50.2)	81.8 (73.0 to 88.3)	67.8 (59.7 to 75.0)
	SQ+	30.9 (21.6 to 42.2)	87.9 (81.5 to 92.3)	65.0 (56.5 to 72.6)
Anxiety[c]	SQ-	19.6 (11.4 to 31.5)	93.9 (91.5 to 95.8)	83.7 (77.9 to 88.2)
	SQ+	27.9 (16.1 to 43.7)	92.5 (88.3 to 95.3)	80.6 (73.5 to 86.2)
Depression, including dysthymia[d]	SQ-	40.6 (31.0 to 51.0)	94.7 (89.4 to 97.4)	87.3 (82.6 to 90.9)
	SQ+	50.0 (36.7 to 63.3)	95.0 (91.0 to 97.2)	89.1 (84.4 to 92.5)
Depression, excluding dysthymia[e]	SQ-	45.8 (34.4 to 57.5)	94.1 (89.1 to 96.9)	89.0 (84.4 to 92.3)
	SQ+	59.8 (48.4 to 70.1)	94.5 (90.8 to 96.8)	91.1 (87.4 to 93.7)
Alcohol abuse[f]	SQ-	25.2 (9.5 to 52.0)	98.9 (97.7 to 99.5)	97.1 (95.4 to 98.2)
	SQ+	60.3 (34.3 to 81.6)	97.9 (96.5 to 98.7)	97.1 (95.5 to 98.2)

[a]Missing GP data on 2 (0.3%) interviewed patients; [b]Somatoform disorders (F44, F45, F48.0); [c]Anxiety, exclusive specific phobias (F40.2) and somatoform disorders (F44, F45, F48.0); [d]Mood disorders, exclusive F30, F31.0–31.2, F31.7–31.9, F33.4; [e]Mood disorders, exclusive F30, F31.0–31.2, F31.7–31.9, F33.4 and 'dysthymia' (F34, F38); [f]Mental and behavioural disorders owing to use of alcohol (F10.0–F10.6).

Table 4. Outcome from screening disclosure on GPs' recognition of various psychiatric disorders. n *= 701[a], weighted data = 1785.*

ICD-10 diagnosis	Absolute benefit increase (ABI)[g] (95% CI)	Relative benefit increase (RBI)[h] (95% CI)	Relative odds ratio (ROR)[i] (95% CI)	Number needed to test (NNT)[j] (95% CI)
1. Somatoform disorders[b]	-5.9 (-22.3 to 10.6)	-15.9 (-48.2 to 36.3)	1.2 (0.4 to 3.3)	-
2. Anxiety disorder[c]	8.3 (-8.9 to 25.5)	42.4 (-30.3 to 191.1)	1.3 (0.4 to 4.0)	12 (-)
3. Depression, including dysthymia[d]	9.4 (-7.6 to 26.3)	23.1 (-14.9 to 78.0)	1.5 (0.5 to 5.0)	11 (-)
4. Depression, excluding dysthymia[e]	14.0 (-2.1 to 30.1)	30.6 (-4.7 to 79.0)	1.9 (0.6 to 5.7)	7 (-)
5. Alcohol abuse[f]	35.2 (2.0 to 68.4)	139.7 (-7.9 to 523.9)	2.3 (0.4 to 13.6)	3 (1 to 29)
Any diagnosis of 1, 2, 3 or 5	-4.5 (-18.0 to 9.1)	-11.1 (-37.7 to 27.0)	0.7 (0.2 to 2.2)	-
Any diagnosis of 2, 3 or 5	7.2 (-7.6 to 22.1)	22.9 (-18.8 to 86.0)	2.8 (1.1 to 7.3)	14 (-)

[a]Missing GP data on 2 (0.3%) interviewed patients; [b]Somatoform disorders (F44, F45, F48.0); [c]Anxiety, exclusive specific phobias (F40.2) and somatoform disorders (F44, F45, F48.0); [d]Mood disorders, exclusive F30, F31.0–31.2, F31.7–31.9, F33.4; [e]Mood disorders, exclusive F30, F31.0–31.2, F31.7–31.9, F33.4 and 'dysthymia' (F34, F38); [f]Mental and behavioural disorders owing to use of alcohol (F10.01–F10.6); [g]ABI = SE (SQ+) – SE(SQ-) where SE = GP sensitivity and SQ = screening questionnaire blinding or disclosure (- or +); [h]RBI = (SE(SQ+) – SE(SQ-)) / SE(SQ-); [i]ROR = OR(SQ+) /OR(SQ-) where OR = true positive∗true negative / false positive∗false negative; [j]NNT = 1/ABI.

may potentially be diluted by: the lack of a 'run-in period' for the screening procedure, a potential Hawthorn bias as a result of the focus on recognition, a cross-contamination of the GPs' diagnostic approaches, a sensitisation of participating patients regardless of randomisation status, better communication skills and psychotherapeutic training status of participating than non-participating GPs, the introduction of an educational programme mainly focusing on somatisation, and the slightly younger age among participating than among non-participating patients, given the observed rise in psychiatric morbidity with age. (Toft *et al*, unpublished data, 2002)

Our screening instrument was evaluated on an appropriate spectrum of patients (i.e. patients consulting their GP for a new complaint). All questionnaires had previously been tested in different settings. Even so, questions may be raised concerning the extent to which the ICD-10 can be regarded as a reference or 'gold standard' in general practice.[16,17] This may especially be the case when trying to make the concept of somatisation operational in primary care,[18] as the term 'medically unexplained physical symptoms' is still causing confusion.[19] In order to reduce the lack of specificity, we chose to validate the GPs' perception of the problem presented (genuinely physical or non-physical) against any ICD-10 diagnosis of somatoform disorders. This approach will tend to underestimate the GPs' recognition of somatoform disorders in cases where the somatising patient presents a genuinely physical condition as the main problem.

Surprisingly, secondary analyses showed evidence of systematic selection bias for SCAN interview among patients with low SQ scores. Patients having a somatoform or anxiety disorder more often accepted interview when SQs had been disclosed (20/47 [43%]) than when SQs were blinded (8/37 [22%], $P = 0.04$). This potential bias could account for the observed differences in the GPs' recognition rates among low-scoring patients (Table 2). We found no way to correct for this. On the other hand, no evidence of selection bias was found concerning randomisation status among the high-scoring patients interviewed.

Moreover, one still has to assess the operational characteristics of the SQ. We reported a PPV of 5% for patients screened negative on Whiteley-7 to the GPs. Using the present data, presuming there was no selection bias for interview, a PPV of 30% would be more appropriate.

Comparison with the existing literature

To our knowledge this study differs from previous randomised, controlled trials on psychiatric screening by giving information on a range of psychiatric disorders. A sys-

Table 5. Outcome from screening disclosure on GPs' recognition of various psychiatric disorders among patients with high scores on the screening questionnaire. n = *617[a], unweighted data.*

ICD-10 diagnosis	Absolute benefit increase (ABI)[g] (95% CI)	Relative benefit increase (RBI)[h] (95% CI)	Relative odds ratio (ROR)[i] (95% CI)	Number needed to test (NNT)[j] (95% CI)
1. Somatoform disorders[b]	11.9 (0.7 to 23.1)	34.3 (1.1 to 78.3)	2.2 (1.1 to 4.3)	8 (4 to 141)
2. Anxiety disorders[c]	16.0 (0.9 to 31.0)	63.0 (0.7 to 163.9)	2.0 (0.8 to 4.7)	6 (3 to 108)
3. Depression, including dysthymia[d]	14.9 (0.5 to 29.3)	36.6 (0.3 to 86.2)	1.5 (0.6 to 3.6)	7 (3 to 200)
4. Depression, excluding dysthymia[e]	14.0 (-2.0 to 30.0)	30.7 (-4.6 to 78.9)	1.5 (0.6 to 3.7)	7 (-)
5. Alcohol abuse[f]	35.0 (2.4 to 67.6)	140.0 (-6.6 to 516.8)	2.0 (0.3 to 11.9)	3 (1 to 42)
Any diagnosis of 1, 2, 3 or 5	15.6 (6.2 to 25.0)	37.2 (12.7 to 67.1)	2.5 (1.2 to 5.3)	6 (4 to 16)
Any diagnosis of 2, 3 or 5	18.8 (7.2 to 30.3)	52.0 (16.5 to 98.4)	3.1 (1.5 to 6.5)	5 (3 to 14)

[a]Missing GP data on 2 (0.3%) interviewed patients. [b]Somatoform disorders (F44, F45, F48.0). [c]Anxiety, exclusive specific phobias (F40.2) and somatoform disorders (F44, F45, F48.0). [d]Mood disorders, exclusive F30, F31.0–31.2, F31.7–31.9, F33.4. [e]Mood disorders, exclusive F30, F31.0–31.2, F31.7–31.9, F33.4 and 'dysthymia' (F34, F38). [f]Mental and behavioural disorders owing to use of alcohol (F10.01–F10.6). [g]ABI = (SE(SQ+) – SE(SQ-) where SE = GP sensitivity and SQ = Screening questionnaire. [h]RBI = (SE(SQ+) – SE(SQ-)) / SE(SQ-). [i]ROR = OR(SQ+) /OR(SQ-) where OR = true positive∗true negative / false positive∗false negative. [j]NNT = 1/ABI.

tematic review of literature by Gilbody *et al*[3] suggests that screening for depression and anxiety may benefit a subgroup of patients with high SQ scores. Our findings suggest that screening for a wider range of psychiatric disorders, including somatisation and alcohol problems, shows similar results. As proposed by Gilbody *et al*, we chose to provide the GPs with PPVs on ratings, thus letting them decide themselves whether to validate proposed diagnoses or not. In this study, 58% of the patients were screened positive on the SQ (dichotomised cut-points on SCL-SOM >3, or Whiteley-7 >0, or SCL-8 >0, or CAGE >0). Hereof 64% (95% CI = 58 to 70) met the ICD-10 criteria of a specific psychiatric diagnosis. Compared to reports on PRIME-MD[5], 81% of the patients were screened positive and 47% of these had a specific psychiatric diagnosis. Using our approach, fewer false-positive results had to be ruled out. However, the GPs would probably have benefited more from the SQ by improving their diagnostic evaluation; for example, by using a diagnostic interview guide.

The implications for future research or clinical practice

Our study demonstrated limited usefulness from routine screening for common psychiatric disorders. However, findings suggest that the SQ may be useful for case-finding among a subgroup of patients with high SQ scores. The development of feasible strategies; for example, a few initial key questions that spot patients at increased risk of having high SQ scores, may prove useful as criteria for case-finding. The significance of having low SQ scores remains unclear and awaits further validation of the instrument.

References

1. Fink P, Sorensen L, Engberg M, *et al*. Somatization in primary care. Prevalence, health care utilization, and general practitioners recognition. *Psychosomatics* 1999; **40(4):** 330-338.
2. Üstün TB, Sartorius N (eds). *Mental illness in general health care: an international study.* Chichester, New York, Brisbane, Toronto, Singapore: John Wiley & Sons, 1996.
3. Gilbody SM, House A, Sheldon TA. Routinely administered questionnaires for depression and anxiety: systematic review. *BMJ* 2001; **322:** 406-409.
4. Staab JP, Datto CJ, Weinrieb RM, *et al*. Detection and diagnosis of psychiatric disorders in primary medical care settings. *Med Clin North Am* 2001; **85(3):** 579-596.
5. Spitzer RL, Williams JB, Kroenke K, *et al*. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA* 1994; **272(22):** 1749-1756.
6. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. *JAMA* 1999; **282(18):** 1737-1744.
7. Fink P, Rosendal M, Toft T. Assessment and treatment of functional disorders in general practice: The extended reattribution and management model — an advanced educational program for non-psychiatric doctors. *Psychosomatics* 2002; **43(2):** 93-131.
8. Derogatis LR, Cleary PA. Confirmation of the dimensional structure of the SCL-90: a study in construct validation. *J Clin Psychol* 1977; **33(4):** 981-989.
9. Fink P, Ewald H, Jensen J, *et al*. Screening for somatization and hypochondriasis in primary care and neurological in-patients: a seven-item scale for hypochondriasis and somatization. *J Psychosom Res* 1999; **46(3):** 261-273.
10. Fink P, Ørnbøl E, Huyse FJ, *et al*. A brief diagnostic screening instrument for mental disturbances in general medical wards — The SCL-8 Scale. A European multicentre study. *J Psychosom Res* 2003 (in press).
11. Fink P, Jensen J, Borgquist L, *et al*. Psychiatric morbidity in primary public health care. A Nordic multicenter investigation. Part I: Method and prevalence of psychiatric morbidity. *Acta Psychiatr Scand* 1995; **92:** 409-418.
12. Ewing JA. Detecting alcoholism: the CAGE questionnaire. *JAMA* 1984; **252:** 1905-1907.
13. WHO. *SCAN. Schedules for clinical assessment in neuropsychiatry, version 2.1.* Geneva: World Health Organization, Division of Mental Health, 1998.
14. World Health Organization. *The ICD-10 classification of mental and behavioural disorders. Diagnostic criteria for research.* Geneva: WHO, 1993.
15. Dunn G, Pickles A, Tansella M, Vazquez-Barquero JL. Two-phase epidemiological surveys in psychiatric research. *Br J Psychiatry* 1999; **174:** 95-100.
16. Kirmayer LJ, Robbins JM. Three forms of somatization in primary care: prevalence, co-occurrence, and sociodemographic characteristics. *J Nerv Ment Dis* 1991; **179(11):** 647-655.
17. Kroenke K, Spitzer RL, deGruy FV, *et al*. Multisomatoform disorder. An alternative to undifferentiated somatoform disorder for the somatizing patient in primary care. *Arch Gen Psychiatry* 1997; **54(4):** 352-358.
18. Schilte AF, Portegijs PJ, Blankenstein AH, Knottnerus JA. Somatisation in primary care: clinical judgement and standardised measurement compared. *Soc Psychiatry Psychiatr Epidemiol* 2000; **35(6):** 276-282.
19. Peveler R, Kilkenny L, Kinmonth AL. Medically unexplained physical symptoms in primary care: a comparison of self-report screening questionnaires and clinical opinion. *J Psychosom Res* 1997; **42(3):** 245-252.

Acknowledgments

The interdisciplinary Research Programme of the Danish National Research Council [Sundhedsfremme og forebyggelsesforskning] (grant number 9801278) and the Health Service of the Aarhus County (project number 0871) funded this study. We thank the physicians, the secretaries and the patients who participated in this study.