Screening for anxiety disorders

Sensitivity and specificity of the Anxiety Screening Questionnaire (ASQ–15)

HANS-ULRICH WITTCHEN and PATRICE BOYER

Background The paper describes the rationale, sensitivity and specificity of the Anxiety Screening Questionnaire (ASQ), a disorder-specific screening instrument for use in primary care.

Method Two hundred and fifty subjects sampled from psychiatric, primary care settings and the community, participated in a test–retest reliability as well as a procedural validity study, using the M–CIDI with DSM–IV algorithms as a diagnostic yardstick.

Results The ASQ was found to be easy to administer and acceptable and efficient in terms of sensitivity and specificity for generalised anxiety syndromes. The test–retest item reliability was good to excellent with kappa values of 0.6 or above. As compared with the validity standard, the DSM–IV/CIDI diagnoses caseness sensitivity was generally high (above 82%) for all diagnostic domains covered, whereas the specificity was only high for DSM–IV threshold and subthreshold generalised anxiety disorder.

Conclusions These preliminary findings demonstrate the usefulness of this anxiety screening questionnaire, constructed closely following the guidelines of specific diagnostic criteria.

Primary care institutions in most health care systems are 'clearing houses' for a considerable number of patients presenting with vague psychological complaints and specific mental disorders. A recent World Health Organization study on mental illness in primary health care settings, conducted in 15 countries around the world (Üstün & Sartorius, 1995, Sartorius *et al*, 1996) demonstrated that among all primary care attenders 24% present with at least one specific mental disorder and an additional 9% with significant subthreshold syndromes. Among all types of disorders examined anxiety disorders were found to be among the most frequent, with generalised anxiety disorder (GAD) being the second most prevalent single diagnosis in primary care (Table 1).

The ability of primary care doctors to recognise mental disorders and anxiety disorders in particular has been shown to be rather poor. Üstin & Sartorius found that only 48.9% of all patients with specific mental disorders were recognised by their primary care doctors as psychiatric cases, and only a few received any type of intervention. Many factors have been held responsible for such poor recognition rates in primary care: vagueness of patient's report, frequency of ill-defined syndromes, tendency of patients to somatise their emotional complaints in an attempt to attract their physicians' attention, very limited time for the primary care doctor to assess the patient, lack of comprehensive psychopathological and differential diagnostic knowledge, lack of sensitive and specific screening tools for use in routine practice.

In response to this critical situation many attempts have been made to improve primary care doctors' ability to detect and manage mental disorders more appropriately. These attempts range from designing specific training courses for primary care doctors (McGlynn & Metcalf, 1992; Montgomery, 1995; Wittchen, 1995), through widely publicised educational packages for the public, patients and their relatives (Wittchen, 1995) to the development of

Table 1 Prevalence and recognition of mental disorders in primary care settings around the world (Sartorius *et al*, 1996)

ICD–10 disorders	Prevalence	Previously recognised[1]
Any depressive disorder	11.7%	6.4%
Current depressive episode (F32/F33)	10.4%	
Dysthmia (F34)	2.1%	
Subthreshold depressive disorders	6.5%	
Any anxiety disorder	10.2%	5.1%
Current generalised anxiety (F41.1)	7.9%	
Panic disorder (F41.0)	1.1%	
Agoraphobia (F40.0)	1.5%	
Subthreshold anxiety disorder	5.0%	
Mixed anxiety–depressive disorder (F41.2)	1.3%	
Somatisation (F45.0)	2.7%	1.5%
Hypochondriasis (F45.2)	0.8%	
Neurasthenia (F48.0)	5.4%	
Alcohol dependence (F10.2)	2.7%	
Harmful use of alcohol (F10.1)	3.3%	

1. Recognition findings are reported for three general categories of disorders.

https://doi.org/10.1192/S000712500029346X Published online by Cambridge University Press

various types of assessment instruments for administration in routine care. The latter pre-formatted diagnostic screening interviews, which require training such as modifications of the Diagnostic Interview Schedule (Q–DIS–3R; Robins *et al*, 1982), the Composite International Diagnostic Interview (CIDI; World Health Organization, 1991), the Structured Clinical Interview (SCID; Spitzer *et al*, 1994), the PRIME–MD (Spitzer *et al*, 1994), or the computerised Symptom Driven Diagnostic System for Primary Care (SDDS–PC; Olfson *et al*, 1995) have, however, to our knowledge, not yet received wider acceptance in the routine work of primary care physicians. Presumably because of the training required for these instruments and the relative complexity of their administration they are unattractive for use in primary care.

Screening questionnaires and rating scales are more attractive for primary care because they do not require specific training with regard to administration and analysis, are less time-consuming and are inexpensive for the clinician. Various broad-spectrum scales with promising psychometric properties have been developed and evaluated. Probably the most widely used, best evaluated and most efficient scale for determining 'caseness' is the General Health Questionnaire (GHQ; Goldberg & Williams, 1988), available in versions of variable lengths. Other more widely used screening instruments for caseness include the Self Reporting Questionnaire (SRQ; Harding *et al*, 1980; World Health Organization, 1994) and the much longer SCL-90 (Derogatis *et al*, 1973) that however has not been developed primarily as a screening device, as were the GHQ and the SRQ. Among the more syndrome-specific screening instruments, the Center for Epidemiological Studies Depression Scale (CES–D; Roberts & Vernon, 1983), the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1993) and the Beck Depression Inventory (BDI; Beck *et al*, 1961) are among the most frequently cited. All these scales provide dimensional measures for one or several domains of psychopathology, indicating the likelihood for caseness in the respective domains assessed, but none offers sufficiently detailed information about the likelihood of having a specific mental disorder. Thus, although most of these scales are quite sensitive in identifying caseness, none has demonstrated sufficient diagnostic specificity.

With the advent of explicit diagnostic criteria for specific types of anxiety and other mental disorders (such as DSM–III (American Psychiatric Association, 1980) and ICD–10 (World Health Organization, 1991)), and increasing evidence for diagnosis-specific management and treatment guidelines for mental disorders (Kasper & Möller, 1995; Margraf 1996), disorder-specific screening questionnaires have recently been receiving more attention. These types of instruments translate directly the specific diagnostic criteria of disorders, as codified in diagnostic classification systems (DSM–IV or ICD–10), into self-report questions, each covering the respective diagnostic criterion. This 'prototypical' approach has been applied to somatoform disorders (Buchholz *et al*, 1993) and depressive as well as specific anxiety disorders. The Inventory to Diagnose Depression (IDD; Zimmerman & Coryell, 1987) as well as the brief screen for panic disorder (Apfeldorf *et al*, 1994) are two examples of this type of screening scale. However, to our knowledge, none of these scales has been comprehensively evaluated in primary care or found widespread use in these settings.

This paper describes the rationale behind and reports preliminary psychometric properties of the Anxiety Screening Questionnaire (ASQ–15), a simple and diagnosis-specific self-report screening questionnaire for GAD and other anxiety syndromes. The following more specific questions will be addressed:

(a) What is the test–retest reliability of the 15 ASQ items?

(b) How sensitive and how specific is the ASQ in identifying general anxiety syndromes as compared to diagnoses derived from a structured diagnostic interview?

(c) How sensitive and specific is the ASQ in detecting other forms of mental disorders?

METHOD

The development of the ASQ–15 and its psychometric evaluation is part of a larger collaborative study between three sites (Munich, Paris and Freiburg). In each of these centres patients from primary care settings, psychiatric institutions and subjects sampled from the community are used to determine the ASQ test–retest reliability as well as various forms of validity. In all centres the same design and instruments described below are being used. As the data collection and analysis has at this point only been completed in the Munich centre, the following results are based on 250 subjects from this site only.

Selection and description of subjects

The study was conducted with 100 in- and out-patients in three psychiatric and one neurological ward of the Clinical Institute of the Max Planck Institute of Psychiatry in Munich, Germany, as well as with 100 patients from the practices of six collaborating primary care doctors in Munich. A further 50 community residents from a larger prospective epidemiological survey, the EDSP (Wittchen *et al*, 1998*c*), were also included. Overall, 250 subjects aged 16–65 completed the full investigation. Exclusion criteria for the study were: obvious severe mental disorders (acute psychotic disorders, dementia and severe depression), current severe withdrawal syndrome and severe impairment in communication.

Psychiatric and neurological sample

We asked 143 in- and out-patients at our Institute, admitted between June and July 1996, to participate in the study. Of these, 11 refused to participate, 28 did not meet the inclusion criteria and four subjects were discharged unexpectedly, not filling out the retest ASQ. Of the remaining 100, 56 were men and 44 women, with an average age of 39.3 (s.d. 18.2) years.

Primary care setting

Although it was originally planned to randomly select primary care attenders, this proved to be impossible. Instead, research assistants went to each of the participating primary care doctors and asked the physician to assign as many of his patients as possible to the investigation. Overall 194 patients were designated by the physicians and were subsequently approached by the research assistant. Twenty-three did not fulfil the inclusion criteria, 32 refused to participate and 39 were either not able or not willing to come in for the retest evaluation and the validation interview. Forty-six per cent of the final primary care sample were men, 54% women, average age was 39.4 years (s.d. 21.3).

https://doi.org/10.1192/S000712500029346X Published online by Cambridge University Press

Community sample

As part of an ongoing prospective epidemiological study of 14–25 year-old residents in the Munich area, 50 subjects aged 16–24 years were randomly chosen to fill out the ASQ during the project interview and were then, as part of the retest examination, sent the ASQ again by mail. As this epidemiological project includes administration of the diagnostic interview, this procedure was quite convenient because no independent diagnostic interview was necessary.

Table 2 summarises the diagnostic distribution and caseness status according to DSM–IV criteria for each of the three groups. Interviews were conducted by trained clinical psychologists using the computerised version of the Munich–Composite Diagnostic Interview (M–CIDI, for a full description, see below) conducted at the time of initial ASQ administration. Overall 71 subjects received no M–CIDI/DSM–IV diagnosis of mental disorder, with the largest proportion of these coming from the community sample and the primary care setting. The remaining subjects revealed considerable comorbidity with an average of 2.2 diagnoses in the psychiatric sample and 1.8 in the community and primary care sample. The most frequent single diagnosis was depressive disorder, followed by GAD. Eleven subjects were diagnosed as having (mostly partially remitted) psychotic disorder.

Instruments

The ASQ (see Appendix) consists of 15 items (codable with 'yes' and 'no') grouped into three sections. Except for the entry question, all questions were taken from the M–CIDI and adapted for self-report administration. Part I of the ASQ consists of one item where the subject is asked to tick the primary reasons for visiting the doctor or for treatment. This question is intended to inform the clinician as to whether the subject suffers from acute pain, psychological or emotional distress, physical or somatic conditions or any other problems. This information is not used for specific diagnostic purposes. Part II consists of six diagnostic questions, each of which addresses a different diagnosis and is used as a stem question for the respective diagnosis. A stem question is defined as the entry criterion of the operationalised diagnostic criteria for a diagnosis and one that has been shown to be highly sensitive for that diagnosis (Kessler *et al*, 1998). A negative answer to this type of question will make it impossible to fulfil the criteria for a full diagnosis. For example, if a person denies having had a panic attack, it is logically impossible to have panic disorder. Similarly, it is not possible to fulfil the criteria for major depression if there has never been a period of two weeks with depressed mood or a loss of interest. The ASQ uses stem questions of this type for major depressive disorders, panic disorder, social phobia, agoraphobia, post-traumatic stress disorder (PTSD) as well as GAD. The efficiency of these stem questions from the CIDI has been demonstrated in numerous methodological studies of the CIDI (see Wittchen & Pfister, 1997*a*,*b*). If all six stem questions are answered in the negative the subject does not need to fill out the subsequent questions in Part III. Part III of the ASQ aims at an evaluation of DSM–IV and ICD–10 GAD as well as subthreshold anxiety syndromes. This part starts with three questions about the contents of the person's anxiety and worries (criterion A1 of DSM–IV), followed by questions to assess the excessiveness of worrying (criterion A2) and the difficulty to stop and to control worrying (criterion B). After this, specific autonomic, muscular, mental and other symptoms of generalised anxiety syndromes are assessed (criterion C of DSM–IV). It is important to note here that the six DSM–IV items (restless, fatigue, impaired concentration, irritability, muscle tension and sleep disturbances) of which three or more need to be fulfilled to meet criterion C, were supplemented by additional items of the ICD–10 diagnostic criteria for GAD. This addition also allows the diagnosis of ICD–10 GAD. The two final questions aim at an evaluation of associated impairments (criterion E of DSM–IV) as well as its duration and persistence (criterion A3). It should be noted that criterion D of DSM–IV specifying numerous hierarchical diagnostic exclusions is not addressed by this questionnaire.

It is noteworthy that in earlier drafts of the ASQ we omitted these subsequent questions in Part III, whenever the stem question for GAD was denied, irrespective of any other endorsed stem question. Our field experience in a pilot test, however, showed that potentially useful information, such as about severity and associated symptoms relevant for other anxiety disorders (panic, social phobia and agoraphobia) PTSD and mixed anxiety–depressive disorders, was lost if they are omitted.

The analysis of the ASQ (see Appendix) allows for the diagnoses of DSM–IV and ICD–10 GAD, as well as the likelihood of major depressive disorders, panic disorder, agoraphobia, social phobia and PTSD. Whenever the full diagnostic criteria for GAD are missed by just one question a subthreshold diagnosis is assigned.

Table 2 Diagnostic characteristics of the reliability and validity study sample (n=250)

DSM–IV diagnosis	Sample by type of setting			
	Psychiatric (n=100)	Primary care (n=100)	Community (n=50)	Total (n=250)
(No diagnosis)	(12)	(38)	(21)	(71 (28.4%))
Generalised anxiety disorder	32	21	7	60 (24.0%)
Panic disorder	23	8	2	33 (13.2%)
Agoraphobia	21	12	4	37 (14.8%)
Social phobia	23	11	7	41 (16.4%)
Post-traumatic stress disorder	12	4	2	18 (7.2%)
Depression	51	24	12	87 (34.8%)
Psychotic disorder	9	2	0	11 (4.4%)
Somatoform disorder	7	11	8	26 (10.4%)
Other disorders	23	19	11	53 (21.2%)
Number of diagnoses	201	112	53	366
Mean number of diagnoses/person	2.2	1.8	1.8	2.0

https://doi.org/10.1192/S000712500029346X Published online by Cambridge University Press

Validation interview: the Munich–Composite International Diagnostic Interview (M–CIDI)

Psychopathological as well as diagnostic assessments were based on the M–CIDI, an updated version of the World Health Organization (WHO) CIDI version 1.2 supplemented by questions of WHO–CIDI version 2.0 developed to cover DSM–IV and ICD–10 criteria (Wittchen & Pfister, 1997*a*,*b*). The M–CIDI allows for the assessment of symptoms, syndromes and diagnoses of 48 mental disorders (not counting various subtypes of main disorders) along with information about onset, duration, clinical and psychosocial severity. Diagnostic analysis is based on the M–CIDI diagnostic package DSM–IV diagnostic algorithms (Pfister & Wittchen, 1995). Diagnostic findings reported in this paper are based on the M–CIDI DSM–IV algorithms without using the DSM–IV hierarchy rules, unless otherwise stated in the text. The M–CIDI includes numerous features that have been developed and tested in several methodological studies with the CIDI or modifications thereof (Kessler *et al*, 1998). These include: (a) the use of symptom lists and memory aids in a separate response booklet to improve lifetime recall, ease memory search as well as to shorten length of the interviews in the somatisation and anxiety section; (b) the addition of symptom and criteria lists to help the proband answer onset and recency questions, for example, in the alcohol section to assess onset and recency of reported dependence symptoms; (c) the implementation ratings in various sections for the assessment of impairment associated with core syndromes; (d) the specific rating of key syndromes for their first, worst and most recent occurrence, with additional questions to allow derivation of pure cross-sectional measures; (e) the incorporation of separate current and lifetime ratings for the degree of impairment in various social roles (work, school, leisure time, partner, etc.); (f) the addition of more open-ended questions describing the person's problems, thereby allowing the clinical editor to judge the appropriateness of the CIDI ratings; (g) the withdrawal of the symptom-specific probe questions of the original CIDI in favour of syndrome based codings; and (h) the deletion of many of the CIDI's skip rules in several diagnostic sections to allow for the study of subthreshold conditions (i.e. mixed anxiety–depression disorders and brief recurrent syndromes) as well as to improve the CIDI's ability to measure more subtle changes in diagnostic status. It is important to note that we did not use the M–CIDI diagnostic section for psychotic disorders in the baseline interview, but only used clinical ratings from the Brief Psychiatric Rating Scale.

The reliability and procedural validity of the CIDI has been established in several studies (Wittchen, 1994). The M–CIDI was additionally tested for test–retest reliability in a two-centre trial of out-patients, as well as a smaller pilot test prior to the beginning of the main study in 14–24-year-olds sampled from the general population (Lachner *et al*, 1998; Wittchen *et al*, 1998*a*). Procedural clinical diagnostic validity was examined against clinician ratings using a DSM–IV and ICD–10 checklist diagnoses (Wittchen *et al*, 1998*a*). The mean duration for completing the computerised M–CIDI, including questionnaires, was 77 minutes. After pilot testing we decided to use the computer-assisted (CAPI) version of the M–CIDI, in order to reduce the length of interview administration, avoid interviewer coding, skip rules and probe question errors, as well as to reduce interviewer variance in formulating questions. The use of the CAPI interview avoided costly key-punching and lengthy data-cleaning procedures, allowing more economic and efficient handling of the data and subsequent data analysis.

Study design

The design allows for an evaluation of the ASQ test–retest reliability as well as its validity in comparison with standardised M–CIDI DSM–IV diagnoses. After signing the informed consent form the subjects fill out the ASQ for the first time (test administration). The M–CIDI was administered immediately afterwards (validation). One to three days later, the ASQ was filled out for the second time to measure the test–retest raliability. For the 26 subjects for whom the three-day time interval was not feasible, questionnaires were returned in these cases four and 11 days later. This design will allow us to evaluate the reliability and validity of the ASQ in comparison with the M–CIDI diagnoses in three strata across all centres.

Analysis

For measuring the test–retest reliability kappa coefficients were calculated for each item as well as for GAD diagnoses. Validity was examined by using the ASQ criteria for DSM–IV threshold and subthreshold GAD and comparing the result with the diagnostic output for these two diagnoses from the M–CIDI, which was used as the 'validity standard' in this study. Subthreshold was defined as falling short of just one DSM–IV criteria A(1) or A(2), B, C or E. In addition, sensitivity (proportion of positive ASQ cases among CIDI cases), specificity (proportion of ASQ non-cases among non CIDI cases), positive (PPV: proportion of CIDI cases among positive ASQ cases) and negative predictive value (NPV: proportion of negative ASQ cases among CIDI non-cases) were determined. Further sensitivity and specificity was also determined for those initial stem questions covering other diagnoses, by comparing the respective ASQ-stem question response to the respective M–CIDI/DSM–IV diagnoses.

RESULTS

Feasibility and acceptance

The mean time for the administration of the ASQ–15 is 3.5 minutes (range 1.2 to 8.3 minutes). Only older, depressed subjects as well as GAD threshold and subthreshold cases need more than 3.5 minutes to complete the questionnaire. Six respondents complained that some of the questions were complicated, long and hard to understand (questions 4 and 6). The vast majority of respondents had no problems with filling out the questionnaire. Regarding the diagnostic analysis two out of the six participating primary care doctors complained about the complicated analysis of the ASQ, because it requires an item-by-item review not allowing the simple addition of all 'yes' responses.

Test–retest reliability

Based on 241 complete test–retest protocols (nine subjects did not return the retest form) the findings shown in Figure 1 suggest that the response patterns in the ASQ questions are stable between the test and the retest administration. Questions with lower kappa values were 'primary reason for contact' (question 1) with κ=0.64, reasons for worrying (question 10, κ=0.64), terrible event (question 6: κ=0.69) and 'difficult to stop the worrying' (question 12: κ=0.68).

https://doi.org/10.1192/S000712500029346X Published online by Cambridge University Press

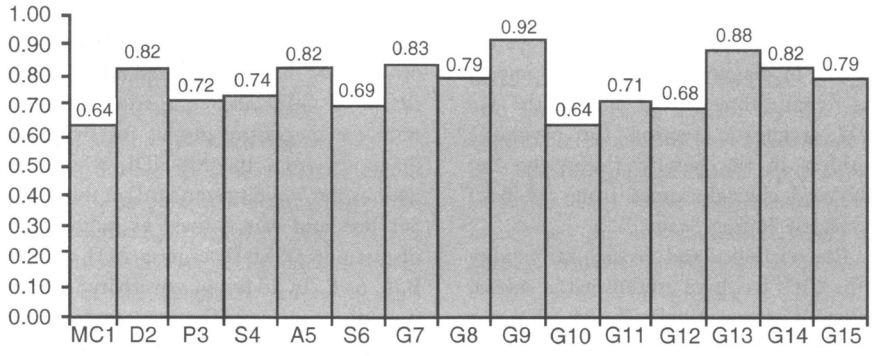

Fig. 1 Test–retest kappa values for each ASQ item. MC, main complaint; D, depression; P, panic; S, social phobia; A, agoraphobia; G, generalised anxiety disorder.

The resulting diagnostic test–retest reliability for ASQ/GAD diagnoses was high with κ=0.72 for threshold GAD and κ=0.70 for subthreshold GAD.

Validity: Comparing the findings of the ASQ–15 administration (test) with the CIDI (validity standard)

In examining the performance of the ASQ, the most critical aspect is how well it detects true cases with threshold GAD (sensitivity) and correctly identifies subjects with no threshold GAD (specificity). For this analysis we use the diagnostic interview diagnostic finding from the CIDI–DSM–IV diagnostic algorithms as the yardstick against which we measure the ASQ performance. Table 3 presents a cross-tabulation as well as the resulting kappa value – as a measure of overall agreement with the sensitivity and specificity findings. Further, we report PPV to examine what proportion with a positive ASQ finding were confirmed by the CIDI and NPV (proportion of non-ASQ cases among non-CIDI cases). Table 3 indicates high overall agreement, resulting in an overall kappa value of 0.88, a high sensitivity of 93.3% and a high specificity of 96.3%. Only seven out of 63 ASQ/GAD cases were not confirmed by the M–CIDI resulting in a slightly lower PPV of 88.9% and in only four cases the M–CIDI assigned a threshold GAD diagnosis not found in the ASQ (NPV: 97.9%).

Subthreshold GAD refers to people who fall short by one diagnostic criterion for the full diagnosis, either not meeting the time, duration or symptom criteria. Table 4 indicates similar high agreement with a kappa value of 0.88 with only slightly diminished PPV and sensitivity.

Sensitivity and specificity of diagnosis-specific ASQ stem questions for other DSM–IV disorders

Extending the analysis to the secondary goal of the ASQ, namely also being able to screen efficiently for other types of disorders, such as panic disorder, agoraphobia, social phobia, depressive and stress-related disorders, Table 5 indicates that the respective stem questions have high sensitivity but only moderate to low specificity. Sensitivity and specificity were both high for depressive and panic disorder only.

Table 3 ASQ–15 threshold GAD caseness against M–CIDI threshold GAD/DSM–IV diagnosis: kappa, negative (NPV) and positive predictive value (PPV), sensitivity and specificity

		ASQ–GAD caseness		
		Yes	No	Total
M–CIDI diagnosis of GAD	Yes	56	4	60
	No	7	183	190
	Total	63	187	250

Kappa: 0.881; PPV: 88.9%; NPV: 97.9%; sensitivity: 93.3%; specificity: 96.3%.

Table 4 ASQ–15 threshold GAD caseness against M–CIDI subthreshold GAD diagnosis[1]: kappa, negative (NPV) and positive predictive value (PPV), sensitivity and specificity

		ASQ–GAD subthreshold		
		Yes	No	Total
M–CIDI subthreshold diagnosis of GAD	Yes	39	3	42
	No	5	203	208
	Total	44	206	250

1. Subthreshold definition: falling short of just one criterion in DSM–IV. Kappa: 0.888; PPV: 88.6%; NPV: 98.5%; sensitivity: 92.8%; specificity: 97.6%.

Table 5 Sensitivity and specificity of the ASQ–15 other marker questions

M–CIDI diagnoses[1]	ASQ–15 stem question findings		
	No. with CIDI diagnosis	Sensitivity (%)	Specificity (%)
Panic disorder (Q.3)	33	95	62
Agoraphobia (Q.5)	37	88	58
Social phobia (Q.4)	41	89	51
Stress/PTSD (Q.6)	18	82	49
Depression (Q.2)	87	98	68
Other disorders (Qs.2,7 and 13)	26	92	74

1. Number in brackets indicates the corresponding ASQ stem question.

https://doi.org/10.1192/S000712500029346X Published online by Cambridge University Press

DISCUSSION

Before discussing the findings of this study in more detail several limitations should be mentioned. (a) The findings reported are based on the preliminary analysis of only one of the three centres involved, which has only 250 subjects up to now. (b) At this preliminary stage, the psychometric properties of the scale are not fully investigated; a major criticism in this respect is the current lack of concurrent validity data with other screening questionnaires. (c) The current validation effort has focused on a comparison of the ASQ–15 stem question and its GAD threshold and subthreshold criteria with the diagnostic findings of M–CIDI, administered right after the subject has filled out the ASQ–15. This design is first most likely to give upper-bounds estimates of the 'true' negative and positive predictive value. Further, since in the validation of such instruments one is most interested in using a 'gold standard', one could also be sceptical in light of the psychometric properties of this instrument for GAD (Wittchen, 1994; Wittchen *et al*, 1996), whether the CIDI is such an optimal tool for diagnosing generalised anxiety syndromes. There is clearly the need to include other diagnostic validation standards (clinical diagnosis and the Structured Clinical Interview for DSM–IV (First *et al*, 1996) are being used in the other centres in the final analyses of the full data set. (d) Finally, it also needs to be mentioned that the statistical procedures used are as yet incomplete. In the final presentation of the full data set, the analysis will include relative (or receiver) operating characteristic curves (Hanley & McNeil, 1982) of the ASQ as well as separate analyses for each setting.

Taking these limitations into account, the ASQ–15, primarily designed as a short and easy to use diagnosis-specific screening questionnaire for generalised anxiety syndromes, was found to be feasible and acceptable for all three groups. Irrespective of type of setting the questionnaire was found to be acceptable to all subjects approached and took only 3–4 minutes to complete. Given the surprisingly high and consistent agreement rates in community respondents as well as in patients in primary care and psychiatric settings a more detailed analysis by setting was not informative. The test–retest reliability of the ASQ–15 items was generally high with kappa values well above 0.6 for each ASQ item. This indicates that within a time span of a week subjects answer the questions consistently in two independent assessments.

There is also excellent agreement of the ASQ-15 diagnostic findings with the results of the computerised threshold and subthreshold CIDI-algorithms for GAD. The kappa values of above 0.8, as well as high NPV and PPV, sensitivity and specificity suggests that the ASQ does very well in discriminating true cases with generalised anxiety syndromes from cases without such diagnoses. An unexpected finding in this respect was that there was no remarkable difference between NPV and PPV. Good screening instruments should have an optimal balance between the two measures but high sensitivity is the most important concern. Showing high PPV and NPV at the same time, the ASQ compares favourably to other screening scales, such as the GHQ (Goldberg & Williams, 1988), the SRQ (World Health Organization, 1994) or to syndrome-specific dimensional scales such as the HADS (Zigmond & Snaith, 1983), that were all shown to have a lower diagnosis-specific PPV. Thus, the present ASQ-15 findings seem to promise an attractive alternative. Given the higher differential effectiveness of specific treatment strategies (Kasper & Möller, 1995; Margraf, 1996) for various kinds of anxiety disorders, the ASQ–15 might have significant advantages over other screening scales, for example when efficient screening for generalised anxiety syndromes in intervention trials is of importance.

It is also noteworthy that even the few initial stem questions for current major depression, as well as other anxiety and stress-related disorders, worked quite well, with high sensitivity estimates of 82 to 95%. The high sensitivity found for these stem questions confirms earlier findings, in which these questions were presented not by a questionnaire but by a trained interviewer (Kessler *et al*, 1998). However, as expected for such a one-item screening the specificity of these stem questions was found to be only moderate, with values ranging from 68% (major depression) to a low of 49% for PTSD. Nevertheless, given their high sensitivity the initial stem questions probably increase the usefulness and attractiveness of the ASQ–15 especially in primary care, where sensitive, easy to use and quick screening devices are the first choice.

REFERENCES

American Psychiatric Association (1980) *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn) (DSM–III). Washington, DC: APA.

— **(1994)** *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM–IV). Washington, DC: APA.

Apfeldorf, W. J., Shear, M. K., Leon, A. C., *et al* (1994) A brief screen for panic disorder. *Journal of Anxiety Disorders*, **8**, 71–78.

Beck, A. T., Ward, C. H., Mendelson, M., *et al* (1961) An inventory for measuring depression. *Archives of General Psychiatry*, **4**, 53–63.

Buchholz, K. K., Dinwiddie, S. H., Reich, T., *et al* (1993) Comparison of screening proposals for somatization disorder: empirical analyses. *Comprehensive Psychiatry*, **34**, 59–64.

Derogatis, L. R., Lipman, R. S. & Covi, L. (1973) The SCL-90: An outpatient psychiatric rating scale. Deutsche Bearbeitung CIPS. *Psychopharmacological Bulletin*, **9**, 13–28.

First, M. B., Gibbon, M., Spitzer, R. L., *et al* (1996) *Users guide for the Structured Clinical Interview for DSM–IV Axis I Disorders – Research Version (SCID–I, Version 2.0)*. New York: Biometrics Research, New York State Psychiatric Institute.

Goldberg, D. & Williams, P. (1988) *A User's Guide to the General Health Questionnaire*. Windsor: NFER-Nelson.

Hanley, J. A. & McNeil, B. J. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **143**, 29–36.

Harding, T. W., DeArango, M. V., Baltazar, L., *et al* (1980) Mental disorders in primary health care: a study of their frequency and diagnosis in four developing countries. *Psychological Medicine*, **10**, 231–241.

Kasper, S. & Möller, H.-J. (eds) (1995) *Angst- un Panikerkrankungen*. Jena: Fischer.

Kessler, R. C., Wittchen, H.-U., Abelson, J. M., *et al* (1998) Methodological studies of the Composite International Diagnostic Interview (CIDI) in the U.S. National Comorbidity Survey. *International Journal of Methods in Psychiatric Research*, **6**, 42–63.

Lachner, G., Wittchen, H.-U., Perkonigg, A., *et al* (1998) Structure, content and reliability of the Munich-Composite International Diagnostic Interview (M-CIDI). Substance use sections. *European Addiction Research*, **4**, 28–41.

Margraf, J. (ed.) (1996) *Lehrbuch der Verhaltenstherapie* (Band 1). Berlin: Springer.

McGlynn, T. J. & Metcalf, H. L. (eds) (1992) *Diagnostik und Therapie von Angststörungen. Ein Handbuch für den Arzt. Übersetzung von: Diagnosis and Treatment of Anxiety Disorders: A Physician's Handbook*. Munich: Eigendruck.

Montgomery, S. A. (ed.) (1995) *Social Phobia: A Clinical Review*. New York: World Psychiatric Association Social Phobia Task Force.

Olfson, M., Leon, M. C., Broadhead, W. E., *et al* (1995) The SDDS-PC. A diagnostic aid for multiple mental disorders in primary care. *Psychopharmacology Bulletin*, **31**, 415–420.

Pfister, H. & Wittchen, H.-U. (1995) *M–CIDI Compueterprogram*. München: Max-Planck-Institut für Psychiatrie, Kinisches Institut.

Roberts, R. E. & Vernon, S. W. (1983) The center for epidemiologic studies depression scale: its use in a community sample. *American Journal of Psychiatry*, **140**, 41–46.

https://doi.org/10.1192/S000712500029346X Published online by Cambridge University Press

APPENDIX – ASQ QUESTIONS AND HOW THEY MATCH DIAGNOSTIC DOMAINS

ASQ Items and Diagnostic Stem Questions	Intent of ASQ Question
1. What is the primary reason for your coming here (tick all that are appropriate)? – pain complaints – psychological and emotional problems – physical complaints/illness, specify: . . . . . . – other reason, specify:	Non-diagnostic entry question Can be used to clarify other problems
2. During the past weeks, have you been suffering from feeling sad, depressed or losing energy for most of the time?	Stem question from M–CIDI for major depressive disorders
3. In the past two weeks, did you experience attacks of anxiety, when all of a sudden you felt frightened, nervous or very uneasy?	Stem question (M–CIDI) for panic syndrome and disorder
4. During the past weeks, have you been bothered by unreasonably strong fears in social situations, like talking to others, doing something in front of others, or being the centre of attention?	Stem question (M–CIDI) for social phobia
5. During the past weeks, have you been bothered by unreasonably strong fears of using public transportation, being in a shop, standing in line, or being in public places?	Shortened stem question (M–CIDI) for agorophobia
6. In the past couple of months, did you experience some unusually terrible or upsetting event or situation or suffer from the after-effects of such an event?	Stem question for DSM/PTSD and acute stress (ICD–10)
7. During the past four weeks, have you been bothered by **feeling worried, tense or anxious most of the time**? ***IF ITEMS 2 to 7 ALL DENIED, SKIP!***	Stem question (M–CIDI) for DSM and ICD GAD
8. Were you anxious and worried about things in your everyday life, like household, work, family, partner, children?	M–CIDI (criterion AI GAD of DSM–IV)
9. Were you worrying about your physical health or somatic illnesses?	(supplementary ICD GAD question)
10. Were you worrying about other things?	(supplementary question differential diagnosis)
11. Did you worry much more than other people in your situation would do?	M–CIDI/DSM–IV criterion A2 (excessive)
12. Do you find it difficult to stop worrying, although you tried hard?	M–CIDI/DSM–IV criterion B M–CIDI/DSM–IV criterion C (at least 3 out of 6): 1
13. When you were feeling worried or anxious, did you feel frequently – restless, frightened or keyed up? – were easily tired and worn out? – did you have difficulty concentrating? – were you nervous and irritable? – did you feel tense, or bothered by aching muscles? – did you have trouble staying or falling asleep? – was your heart pounding or racing frequently? – did you feel trembly or shaky? – did you sweat a lot? – did you have difficulty breathing? – very upset, because of your worrying?	2 3 4 5 6 The remaining items are supplements to cover ICD–10
14. Did this feeling worried or anxious interfere a lot with your everyday activities in work, household or the relationship to others?	M–CIDI/DSM–IV criterion E
15. When did this period of worrying start? Was it: – weeks? – months? – or years ago?	M–CIDI/DSM–IV onset and duration criteria (A3)

Robins, L. N., Helzer, J. E., Ratcliff, K. S., *et al* (1982) Validity of the Diagnostic Interview Schedule, version II: DSM–III diagnoses. *Psychological Medicine*, **12**, 855–870.

Sartorius, N., Üstün, T. B., Lecrubier, Y., *et al* (1996) Depression comorbid with anxiety: results from the WHO study on psychological disorders in primary health care. *British Journal of Psychiatry*, **168** (suppl. 30), 29–34.

Spitzer, R. L., Williams, J. B. W., Kroenke, K., *et al* (1994) Utility of a new procedure for diagnosis mental disorders in primary care. The PRIME-MD 1000 Study. *Journal of the American Medical Association*, **272**, 1749–1756.

Üstün, T. B. & Sartorius, N. (eds) (1995) *Mental Illness in General Health Care. An International Study.* Chichester: Wiley.

Wittchen, H.-U. (1994) Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research*, **28**, 57–84.

—— (1995) Short- and longterm effectiveness of a training programme for recognition. *Primary Care Psychiatry*, **1**, 45–52.

——, Beloch, E., Garczynski, E., *et al* (1996) *Münchener Composite International Diagnostic Interview* (M–CIDI, Paper-pencil 2.2, 2/95). München: Max-Planck-Institut für Psychiatrie, Klinisches Institut (Eigendruck).

——, Nelson, G. B. & Lachner, G. (1998*a*) Prevalence of mental disorders and psychosocial impairments in adolescents and young adults. *Psychological Medicine*, **28**, 109–126.

——, Lachner, G., Wunderlich, U., *et al* (1998*b*) Test–retest reliability of the computerized DSM–IV version of the Munich-Composite International Diagnostic Interview (M–CIDI). *Social Psychiatry and Psychiatric Epidemiology*, in press.

——, Perkonigg, A., Lachner, G., & Nelson, C. B. (1998*c*) The early developmental stages of

https://doi.org/10.1192/S000712500029346X Published online by Cambridge University Press

psychopathology study (EDSP) – objectives and design. *European Addiction Research*, **4**, 18–27.

__ & Pfister, H. (1997*a*) *DIA-X-Interview Manual. Istruktionsmanual zur Durchführung von DIA-X-Interviews.* Frankfurt: Swets & Zeitlinger.

__ & __ (eds) (1997*b*) *DIA-X-Interviews: Manual für Screening-Verfahren und Interview: Interviewheft Längsschnittuntersuchung (DIA-X-Lifetime); Ergänzungsheft (DIA-X-Lifetime): Interviewheft Querschnittuntersuchung (DIA-X-12 Monate); Ergänzungsheft (DIA-X-12Monate); PC-Programm zur Durchführung des Interviews (Längs- und Querschnittuntersuchung); Auswertungsprogramm.* Frankfurt: Swets & Zeitlinger.

HANS-ULRICH WITTCHEN, Clinical Psychology and Epidemiology, Max Planck Institute of Psychiatry, München, Germany; PATRICE BOYER, INSERM, Hôpital Ste Anne, CMME, Paris, France

Correspondence: Professor H.-U. Wittchen, Max Planck Institute of Psychiatry, 8004 München, Kraepelinstr. 2, Germany

World Health Organization (1991) *ICD–10, Chapter V, Mental and Behavioural Disorders* (including disorders of psychological development). Diagnostic Criteria for Research, Draft for Field Trials. Geneva: WHO.

__ (1994) *A User's Guide to the Self-reporting Questionnaire* (SRQ). Geneva: WHO.

Zigmond, A. S. & Snaith, R. P. (1983) The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, **67**, 361–370.

Zimmerman, M. & Coryell, W. (1987) The inventory to diagnose depression (IDD): a self-report scale to diagnose major depressive disorder. *Journal of Consulting and Clinical Psychology*, **55**, 55–59.

https://doi.org/10.1192/S000712500029346X Published online by Cambridge University Press