Screening instruments for anxiety and depression in patients with irritable bowel syndrome are ambiguous

Anne Rode Larsen, Anne Line Engsbro & Peter Bytzer

ABSTRACT

INTRODUCTION: Psychiatric disorders are common among patients with irritable bowel syndrome (IBS) which may make a psychiatric evaluation appropriate. The Hospital Anxiety and Depression Scale (HADS) has been recommended internationally for screening. The Common Mental Disorder Questionnaire (CMDQ) is a novel Danish instrument recommended for use in general practice, but it lacks validation. The objective of this study was to test the reliability and the convergent validity of the anxiety and depression subscales of the CMDQ and the HADS in patients with IBS.
MATERIAL AND METHODS: A total of 149 patients with IBS completed the CMDQ and the HADS twice in a year. Data were analysed with respect to internal consistency (Cronbach's alpha), agreement on case identification (κ) and correlation between scores (Spearman's rank correlation coefficient).
RESULTS: All subscales showed satisfactory internal consistency. Cases identified by the CMDQ's anxiety and depressions subscales differed considerably from those identified with the HADS with only fair-moderate agreement, while the correlation between the scores on the CMDQ and the HADS was only moderate for both anxiety and depressions.
CONCLUSION: The subscales for anxiety disorder and depression on the CMDQ showed internal consistency, but only fair to moderate agreement and correlation, which yielded an unsatisfying convergent validity compared to the HADS. More studies in different populations and research on the cut-off values for possible cases could make the CMDQ more useful.
FUNDING: Danish Council for Independent Research; Region Zealand's Health Sciences Research Foundation; Councils for Quality Assurance in Primary Care in Region Zealand.
TRIAL REGISTRATION: NCT00659763.

Irritable bowel syndrome (IBS) is a common disorder characterised by abdominal pain and change in stool habits [1]. Patients with IBS have impaired quality of life, increased sick leave and greater health-care utilisation [2]. High costs are partly related to psychiatric co-morbidity [3, 4].

The prevalence of psychiatric disorders among IBS patients has been reported to range from 54% to 94% [3]. It has therefore been argued that investigation for IBS should include psychiatric evaluation, and the Hospital Anxiety and Depression Scale (HADS) has been recommended internationally for screening purposes [5].

The Common Mental Disorder Questionnaire (CMDQ) was designed for screening for somatoform disorders, mental disorders and alcohol dependence in general practice [6]. It is included in the Guidelines for Anxiety from the Danish College of General Practitioners (DSAM) [7], and it may thus have an impact on the diagnosing of these disorders in primary care in Denmark. However, research on the validity of the CMDQ is limited [6, 8, 9] and the CMDQ lacks internal and external validation.

Our aim was to assess the reliability of the CMDQ and the convergent validity of the anxiety and depression subscales of the CMDQ and the HADS in a population of patients with IBS.

MATERIAL AND METHODS
Study population
This study is part of a randomised trial on diagnostic strategies for suspected IBS in primary care. It was conducted at Køge Hospital, Denmark and in Odense, Denmark. Only data on the Køge population (n = 149) are included in the present analysis. Patients were referred by their general practitioner based on a clinical suspicion of IBS. Eligible patients were 18-50 years and met the Rome-III criteria for IBS [1]. The trial compared a positive diagnostic strategy with a strategy of exclusion. No further interventions or treatment were given. Further details on the study design and major outcomes are reported elsewhere [10].

Questionnaires
Common Mental Disorder Questionnaire
The CMDQ consists of 37 items divided into six subscales and one item on overall health. It includes two subscales for somatoform disorder (SCL-SOM symptom checklist and Whiteley-7 Illness Worry Scale), three subscales for mental disorder (SCL-8 for mental disorder in general, SCL-ANX4 for anxiety disorder and SCL-DEP6 for depression) and one subscale (CAGE) for alcohol abuse and dependence [6]. To compare the CMDQ with the HADS, only the SCL-ANX4 and the SCl-DEP6 were examined.
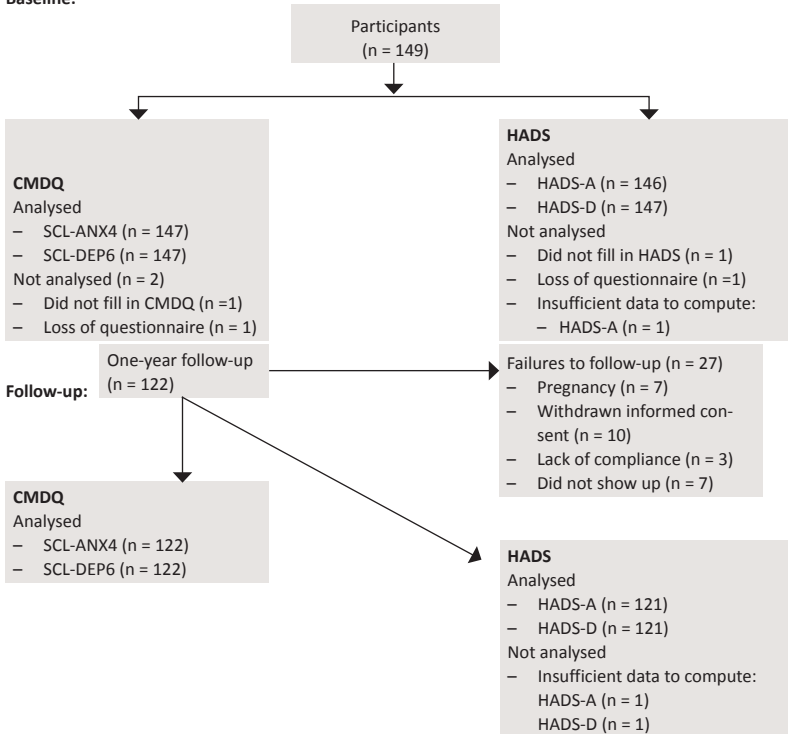
FIGURE 1

Patient trial flow.

CMDQ = Common Mental Disorder Questionnaire; HADS = Hospital Anxiety and Depression Scale; HADS-A = Hospital Anxiety and Depression Scale, anxiety subscale; HADS-D = Hospital Anxiety and Depression Scale, depression subscale; SCL-ANX4 = Common Mental Disorder Questionnaire, anxiety subscale; SCL-DEP6 = Common Mental Disorder Questionnaire, depression subscale.

The SCl-ANX4 consists of four items (item 20-23): "During the last four weeks how much were you bothered by...": 20) "feeling suddenly scared for no reason?", 21) "nervousness or shakiness inside?", 22) "spells of terror or panic?", 23) "you worry too much?".

The SCL-DEP6 consists of six items (item 27-32): "During the last four weeks, how much were you bothered by..": 27) "feeling blue?", 28) "feelings of worthlessness?", 29) "thoughts of ending your life?", 30) "feelings of being trapped or caught?", 31) "feeling lonely?", 32) "blaming yourself for things?".

Answers are rated on five-point Likert-scales ranging 0-4 ("not at all"-"extremely") and are then dichotomised to 0 ("not at all") or 1 ("a little"-"extremely") and summed for each subscale. The sumscore for anxiety ranges 0-4 and for depression 0-6. A sumscore of ≥ 3 is suggested by the authors of CMDQ as a cut-off value for case identification on both subscales [6]. An exact definition of case identification has not been made. The Guidelines for Anxiety from the DSAM [7] defines a score on the SCL-ANX4 of ≥ 2 as abnormal, which was also used in our analyses.

The authors of the CMDQ recommend that unanswered items are set to 0 [6]. We accommodated this only if ≤ 50% of items were missing. Otherwise, the patient was excluded from analyses on the subscale in question.

Hospital Anxiety and Depression Scale
The HADS was developed in 1983 for screening in outpatient departments [11]. It is widely used and has been validated in a broad spectrum of populations, including primary care [12]. It consists of 14 items; seven items forming an anxiety subscale (HADS-A) and seven items comprising a depression subscale (HADS-D). The recall-period is one week. Each item is rated from zero to three, yielding a sumscore for each subscale of 0-21 [11, 12]. A cut-off value for possible cases of ≥ 8 is recommended for both subscales [11-13]. If only one item was missing, unanswered items were handled by using the mean of answered items in the subscale. Otherwise, the patient was excluded from the analyses of the subscale in question. All patients in the trial were asked to fill in both questionnaires at baseline and at follow-up one year later (**Figure 1**).

Statistics
Baseline and follow-up results were compared using Student's t-test for normally distributed data and the Mann-Whitney test for non-normally distributed data. As a measure of the subscales' internal consistency and thereby their reliability, we calculated Cronbach's alpha (α), which should be > 0.70 to be satisfactory [14].

Based on the recommended cut-off values, the proportion of possible cases of anxiety disorder and depression were calculated, and the proportions at baseline and at follow-up were compared using McNemar's test for paired data. For evaluation of the convergent validity of the CMDQ compared with the HADS, dichotomous results, i.e. case identification, were compared by use of κ statistics; and continuous data, i.e. the scores, were compared by the Spearman's rank correlation coefficient (r_s) [14].

κ is usually interpreted as follows: < 0.20 poor agreement, 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 good and 0.81-1.00 very good agreement [15]. For r_s, ± 0.2 is a weak correlation, ± 0.5 a moderate and ± 0.8 a strong correlation [16]. All statistical analyses were performed using SPSS for Windows, version 18.0.

Trial registration: NCT00659763.

RESULTS
The demographic characteristics are summarised in **Table 1**. Patients lost to follow-up did not differ from the entire baseline population.

Mean (standard deviation) scores on the anxiety subscales and median scores (interquartile range) on the depression subscales are shown in **Table 2**. No significant changes in scores were observed from baseline to follow-up.

Reliability

The internal consistency of the subscales on the CMDQ estimated by Cronbach's alpha, was α = 0.82 for SCL-ANX4 and α = 0.87 for SCL-DEP6 at baseline. Similar results were found at follow-up; α = 0.79 for SCL-ANX4 and 0.87 for SCL-DEP6. The HADS showed similar values with α = 0.79-0.83 for HADS-A and α = 0.84-0.87 for HADS-D at baseline and follow-up, respectively.

Case identification and agreement

The proportion of possible cases according to the different subscales and the agreement between the corresponding subscales including κ-values are shown in **Table 3**.

Case identification varied considerably between the CMDQ and the HADS. The agreement between the corresponding subscales for anxiety and depression of the CMDQ and the HADS was fair to moderate (κ = 0.38-0.55). We found no significant change in the proportion of possible cases from baseline to follow-up on any of the subscales (p = 0.09-0.50 for SCL-ANX4; p = 0.19 for HADS-A; p = 0.22 for SCL-DEP6; p = 1.00 for HADS-D).

Correlation between the anxiety and depression subscales on the CMDQ and the HADS

At baseline, the scores on the SCL-ANX4 correlated with the HADS-A (r_s = 0.73; $p < 0.001$). The SCL-DEP6 showed a weaker correlation with the HADS-D (r_s = 0.55; $p < 0.001$).

At follow-up, the correlation was similar for the anxiety subscales (r_s = 0.76; $p < 0.001$), but higher for the depression subscales (r_s = 0.71; $p < 0.001$).

DISCUSSION

We aimed to assess the reliability of the CMDQ and the convergent validity of the anxiety and depression subscales of the CMDQ and the HADS in a population of patients with IBS. The investigated subscales showed a satisfactory internal consistency, but convergent validation showed that the case identification varied considerably between the CMDQ and the HADS, and that the agreement between the two screening instruments was only fair to moderate. The correlation between scores on the two CMDQ's subscales and the HADS was only moderate as far as both anxiety disorder and depression were concerned.

The existing research on the CMDQ's reliability and validity is limited with only three published studies [6, 8,

TABLE 1

Demographic characteristics of the sample.

	Baseline (N = 149)	1-year follow up (N = 122)
Gender		
Females, n (%)	127 (85)	105 (86)
At inclusion		
Age, yrs, mean (SD)	32.4 (8.7)	33.7 (9.1)
Age, yrs, range	18-50	18-50
≤ 35 yrs, n (%)	87 (58)	65 (53)
Ethnicity		
Danish, n (%)	137 (92)	111 (91)

SD = standard deviation.

TABLE 2

Scores on the anxiety and depression scales at baseline and follow-up.

	Baseline	Follow-up	Change, p-value
Anxiety subscales, mean score (± SD)			
SCL-ANX4: score 0-4	1.78 (± 1.25)	1.62 (± 1.31)	0.31
n	147	122	
HADS-A: score 0-21	6.47 (± 3.86)	6.20 (± 4.02)	0.57
n	146	121	
Depression subscales, median score (± IQR)			
SCL-DEP6: score 0-6	2 (± 4.0)	1 (± 4.0)	0.73
n	147	122	
HADS-D: score 0-21	3 (± 5.0)	2 (± 5.0)	0.15
n	147	121	

SD = standard deviation; HADS-A = Hospital Anxiety and Depression Scale, anxiety subscale; HADS-D = Hospital Anxiety and Depression Scale, depression subscale; IQR = interquartile range; SCL-ANX4 = Common Mental Disorder Questionnaire, anxiety subscale; SCL-DEP6 = Common Mental Disorder Questionnaire, depression subscale.

9]. Only one of these studies has been made by other scientists than the authors of the CMDQ. The internal consistency has not been tested, but as a measure of reliability, the scales' uni-dimensionality was found to be satisfactory in a waiting room population in primary care [8]. In the same population, the case-finding abilities of the SCL-ANX4 and the SCL-DEP6 were found to be excellent compared with the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) yielding International Classification of Diseases (ICD)-10 diagnoses as a gold standard [6]. Similar results were found in a population with long-term sick leave with an efficiency, i.e. correct classification, of 76% when using a cut-off value ≥ 3 on both subscales [9]. We did not compare our results on the CMDQ with ICD-10 diagnoses and thus comparison with these earlier findings is not possible.

Criterion validation is part of the testing of a scale's validity. Since the ICD-10 or Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV diagnoses were

▤ **TABLE 3**

Identification of possible cases of anxiety and depression as estimated by the CMDQ and the HADS, and agreement between corresponding subscales. Results from baseline and follow-up.

| | Possible cases, n (%) | | Agreement with HADS-A, cut-off value ≥ 8 | | | |
| | | | proportion of agreement, % | | κ (95% CI) | |
	baseline	follow-up	baseline	follow-up	baseline	follow-up
Anxiety subscales						
SCL-ANX4						
Cut-off value ≥ 2	88 (60)	64 (53)	71	74	0.44 (0.30-0.59)	0.48 (0.33-0.64)
Cut-off value ≥ 3	39 (27)	28 (23)	73	81	0.40 (0.24-0.56)	0.55 (0.37-0.72)
HADS-A						
Cut-off value ≥ 8	58 (40)	38 (31)	–	–	–	–
Depression subscales						
SCL-DEP6						
Cut-off value ≥ 3	56 (38)	45 (37)	73	75	0.38 (0.21-0.54)	0.41 (0.23-0.59)
HADS-D						
Cut-off value ≥ 8	30 (20)	23 (19)	–	–	–	–

HADS-A = Hospital Anxiety and Depression Scale, anxiety subscale; HADS-D = Hospital Anxiety and Depression Scale, depression subscale; SCL-ANX4 = Common Mental Disorder Questionnaire, anxiety subscale; SCL-DEP6 = Common Mental Disorder Questionnaire, depression subscale.

not evaluated in this trial, we lack a gold standard which is a limitation. However, our study was conducted in a homogenous and well-defined population, which makes it suitable for investigating the psychometric properties of the scale. Furthermore, IBS patients are a new population compared to those investigated in previous studies on the CMDQ. Thus, our study contributes to the evaluation of the scales' external validity.

Given the lack of a gold standard, we used the HADS as a criterion measure. This seems reasonable since the HADS has been widely used and extensively validated in several countries and patient populations since 1983. Two systematic reviews and a meta-analysis conclude that the HADS is reliable and valid in both somatic and psychiatric patient settings, both in primary care, in hospital settings and in the general population [12, 13, 17].

Several studies have investigated the prevalence of psychiatric disorders among IBS patients, but comparison is difficult due to large differences in populations and methodology. The prevalence of depression has been estimated to 19-39% [18-20]. Our study showed a similar prevalence using both the SCL-DEP6 (37-38%) and the HADS-D (19-20%). The prevalence of anxiety was 13-36% [18-20] in previous studies. Our findings were consistent with this using both the SCl-ANX4 with a cut-off value ≥ 3 (23-27%) and the HADS-A (31-40%). When using a cut-off value ≥ 2 on the SCL-ANX4, we found a higher prevalence of 53-60%, which could imply more false positives with this cut-off.

A strength of our study is that the same population filled in the questionnaires twice, which allowed us to assess the consistency of the psychometric properties. Furthermore, the response rates were high with 99% of the patients completing the questionnaires at baseline.

We had a failure rate of 18% at follow-up, which is a potential bias. When comparing scores at baseline between the drop-outs and the patients with succesful follow-up, the two populations were not different (p = 0.75 for SCl-ANX4; p = 0.20 for HADS-A; p = 0.89 for SCL-DEP6; p = 0.50 for HADS-D). Furthermore, the demographic characteristics of the sample were unchanged at follow-up. Thus, we think that patients lost to follow-up do not bias our results.

The authors of the CMDQ found that patients with mental disorders have more missing responses than patients without mental disorders [6]. This collides with the recommendation of setting missing items as "0"/"not at all". In our study, the frequency of missing items was very low and when re-analysing data without correcting the missing items, we obtained comparable results. The bias introduced by correction of missing items is thus not significant.

Although the CMDQ and the HADS are theoretically thought to measure the same condition, there are some limitations to their comparability. The recall-periods are different; four weeks on the CMDQ and one week on the HADS. This difference may impact the agreement and the correlation, especially if the underlying disorder is fluctuating. Finally, the different cut-off values recommended for the SCL-ANX4, i.e. ≥ 3 or ≥ 2 [6, 7], contribute to the confusion, and a more thorough information of the definition and the consequenses of the dichotomisation made by the cut-off values could improve the usage.

Neither the CMDQ nor the HADS contain items concerning somatic aspects of anxiety and depression. When using them for screening in patients with gastrointestinal symptoms, this should theoretically decrease confounding, but it could also be a limitation for both questionnaires, since some patients with mental disorders present with somatic symptoms. This somatisation introduces a challenge for physicians, especially with a condition like IBS where there is A considerable interplay between somatic and psychological symptoms. The CMDQ contains a subscale for somatization (SCL-SOM), but an evaluation of this was not part of our study.

At follow-up, patients were asked to rate their current severity of gastrointestinal symptoms compared with baseline on a seven-point scale [10]. Surprisingly, we saw no difference between patients with unchanged or worsened symptoms relative to patients with symptom improvement when we compared their changes in scores from baseline to follow-up on the CMDQ or the HADS. Thus, the development in somatic symptoms did

not seem to influence the emotional symptoms measured by the two questionnaires.

CONCLUSION

The subscales for anxiety disorder and depression of the CMDQ were found to be internally consistent, but the agreement on case identification with the HADS was low and the correlation between the CMDQ and the HADS scores was only moderate, thus yielding an unsatisfactory convergent validity of the CMDQ's anxiety and depression subscales when compared to the HADS. The psychometric properties of the CMDQ and the scale's diagnostic efficacy should be tested in different populations as part of an external validation study. More studies should be conducted to determine appropriate cut-off values on the CMDQ scores, and the guidelines for interpretation of the scores should be clarified.

CORRESPONDENCE: *Anne Rode Larsen*, Medicinsk Afdeling, Køge Sygehus, Lykkebækvej 1, 4600 Køge, Denmark. E-mail: Anne_Rode@hotmail.com

LITERATURE

1. Longstreth GF, Thompson WG, Chey WD et al. Functional bowel disorders. Gastroenterol 2006;130:1480-91.
2. Agarwal N, Spiegel BM. The effect of irritable bowel syndrome on health-related quality of life and health care expenditures. Gastroenterol Clin North Am 2011;40:11-9.
3. Whitehead WE, Palsson O, Jones KR. Systematic review of the comorbidity of irritable bowel syndrome with other disorders: what are the causes and implications? Gastroenterol 2002;122:1140-56.
4. Quigley EM, Bytzer P, Jones R et al. Irritable bowel syndrome: the burden and unmet needs in Europe. Dig Liver Dis 2006;38:717-23.
5. Spiller R, Aziz Q, Creed F et al. Guidelines on the irritable bowel syndrome: mechanisms and practical management. Gut 2007;56:1770-98.
6. Christensen KS. A brief case-finding questionnaire for common mental disorders: the CMDQ. Fam Pract 2005;22:448-57.
7. Christensen KS, Nielsen L, Rosenberg N et al. Klinisk vejledning for almen praksis: angsttilstande – diagnostik og behandling. Dansk Selskab for Almen Medicin, 2010.
8. Christensen KS, Bech P, Fink P. Measuring mental health by questionnaires in primary care – unidimensionality, responsiveness and compliance. Eur Psy Rev 2010;3:8-12.
9. Søgaard HJ. Choosing screening instrument and cut-point on screening instruments. A comparison of methods. Scand J Publ Health 2009;37:872-80.
10. Begtrup LM, Engsbro AL, Kjeldsen J et al. A positive diagnostic strategy is non-inferior to a strategy of exclusion for patients with irritable bowel syndrome. Clin Gastroenterol Hepatol 2013;11:956-62.
11. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361-70.
12. Bjelland I, Dahl AA, Haug TT et al. The validity of the Hospital Anxiety and Depression Scale. J Psychosom Res 2002;52:69-77.
13. Herrmann C. International experiences with the Hospital Anxiety and Depression Scale - a review of validation data and clinical results. J Psychosom Res 1997;42:17-41.
14. Streiner DL, Norman GR. Health measurement scales – a practical guide to their development and use. Third ed. Oxford University Press, 2003.
15. Altman DG. Practical statistics for medical research. Chapman & Hall/CRC, 1991.
16. Zou KH, Tuncali K, Silverman SG. Correlation and simple linear regression. Radiology 2003;227:617-22.
17. Brennan C, Worrall-Davies A, McMillan D et al. The Hospital Anxiety and Depression Scale: a diagnostic meta-analysis of case-finding ability. J Psychosom Res 2010;69:371-8.
18. Jones R, Latinovic R, Charlton J et al. Physical and psychological co-morbidity in irritable bowel syndrome: a matched cohort study using the General Practice Research Database. Aliment Pharmacol Ther 2006;24:879-86.
19. Ladabaum U, Boyd E, Zhao WK et al. Diagnosis, comorbidities, and management of irritable bowel syndrome in patients in a large health maintenance organization. Clin Gastroenterol Hepatol 2012;10:37-45.
20. Whitehead WE, Palsson OS, Levy RR et al. Comorbidity in irritable bowel syndrome. Am J Gastroenterol 2007;102:2767-76.