

The accuracy of HADS and GHQ-12 in detecting psychiatric morbidity in breast cancer patients

Margörit Rita KRESPI BOOTHBY^{1,2}, Jonathan HILL³, Christopher HOLCOMBE⁴, Louise CLARK², Jean FISHER⁴, Peter SALMON²

Abstract

Objective: Psychological problems should be identified in breast cancer patients proactively if doctors and nurses are to help them cope with the challenges imposed by their illness. Screening is one possible way to identify emotional problems proactively. Self-report questionnaires can be useful alternatives to carrying out psychiatric interviews during screening, because interviewing a large number of patients can be impractical due to limited resources. Two such measures are the Hospital Anxiety and Depression Scale (HADS) and the General Health Questionnaire-12 (GHQ-12).

Method: The present study aimed to compare the performance of the GHQ-12, and the HADS Unitary Scale and its subscales to that of the Schedule for Affective Disorders and Schizophrenia (SADS) in identifying patients with affective disorders, including DSM major depression and generalized anxiety disorder. The sample consisted of 296 female breast cancer patients who underwent surgery for breast cancer a year previously.

Results: A small number of patients (11%) were identified as having DSM major depression or generalized anxiety disorder based on SADS score. The findings indicate that the optimal thresholds in detecting generalized anxiety disorder and DSM major depression with the HADS anxiety and depression subscales were ≥ 8 and ≥ 7 , with 93.3% and 77.3% sensitivity, respectively, and 77.9% and 87.1% specificity, respectively. They also had a 21% and 36% positive predictive value, respectively. Using the HADS Unitary Scale the optimal threshold for detecting affective disorders was ≥ 12 , with 88.9% sensitivity, 80.7% specificity, and a 35% positive predictive value. In detecting affective disorders, the optimal threshold on the GHQ-12 was ≥ 2 , with 77.8% sensitivity and 70.2% specificity. This scale also had a 24% positive predictive value. In detecting generalized anxiety disorder and DSM major depression, the optimal thresholds on the GHQ-12 were ≥ 2 and ≥ 4 with 73.3% and 77.3% sensitivity, respectively, and 67.5% and 82% specificity, respectively. The scale also had 12% and 29% positive predictive values, respectively.

Conclusion: The HADS Unitary Scale and its subscales were effective in identifying affective disorders. They can be used as screening measures in breast cancer patients. The GHQ-12 was less accurate in detecting affective disorders than the HADS, but it can also be used as a screening instrument to detect affective disorders, generalized anxiety disorder, and DSM major depression.

Key Words: Breast cancer, GHQ-12, HADS and SADS

INTRODUCTION

Despite advances in medical technology, breast cancer remains a life-threatening illness and imposes enormous demands and challenges on patients. In particular, breast cancer and its treatment place patients at risk for emotional distress. Emotional distress is a psychological,

cally, socially, and spiritually unpleasant experience, and can potentially interfere with a patients' ability to cope with cancer and its treatment. This includes not only common and normal feelings of helplessness, sadness, and fear, but also serious psychological problems such as clinical depression and anxiety. Among cancer patients clinical depression and anxiety are common.

Received: 18.11.2008 - Accepted: 14.07.2009

^{1,2}Psychologist/Assist. Prof., Department of Psychology, Istanbul Bilim University and Division of Clinical Psychology, University of Liverpool, ³MD/Prof., Department of Psychiatry, University of Manchester, ⁴MD/Prof., Breast Unit, Royal Liverpool University Hospital, ²Psychologist, Division of Clinical Psychology, University of Liverpool, ⁴Nurse, Breast Unit, Royal Liverpool University Hospital, ²Psychologist/Prof., Division of Clinical Psychology, University of Liverpool.

This paper was presented as a poster at the British Psychological Society, Division of Health Psychology Conference held in September 2009.

Margörit Rita Krespi Boothby, e-mail: rkrespi@hotmail.com

At some point during their disease journey, a significant percentage (20%-35%) of female breast cancer patients suffer from symptoms of depression and anxiety, irrespective of disease stage and treatment status (Fallowfield et al. 1990). On the other hand, around the time of surgery 33% of patients experience clinical levels of anxiety or depression and 50% of patients experience these problems some time during the first year (Burgess et al. 2005). One year following surgery for breast cancer a significant percentage of patients experience symptoms of anxiety (21%-27%) and depression (14%-21%) (Goldberg et al. 1992; Nosarti et al. 2002). Despite the fact that these problems are common in cancer patients, only a small number are identified and treated (Maguire 1985). It is unfortunate that the remaining patients are left to their own devices to cope with such debilitating psychological problems in addition to the challenges imposed by their cancer and its treatment.

At different stages of cancer clinical depression and anxiety are often related to the nature and intensity of the challenges experienced (White and Macleod 2002). For example, around the time of diagnosis, patients are likely to experience mixed and complex feelings, often related to difficulties in accepting and coping with the illness. These may include shock, anger, fear, and suspicion. Moreover, emotional problems that are experienced during this time are considered to be an important predictor of emotional problems, especially clinical depression, at later stages of cancer. In the treatment stage patients may experience depression and/or anxiety, as they are required to cope with surgery and the side effects of radiotherapy and/or chemotherapy. Following treatment, patients may experience these problems due to preoccupation or fear of recurrence or metastases. Moreover, around this time, although some patients review their life and attach positive meaning to their illness (Salmon et al. 1996), others can experience emotional distress due to, for example, difficulties in personal relationships or personal losses that resulted from cancer and its treatment.

Due to limited resources every depressed cancer patient cannot be assessed and treated by a psychologist or psychiatrist. During the course of routine consultation doctors and nurses can identify these patients' psychological problems and can help them cope with the challenges imposed by cancer (Salmon 2000). This can supplement and complement existing resources. In order to do this, psychological problems should be identified proactively. One possible way to identify emotional problems proactively is to screen patients during routine care.

In clinical settings standardized psychiatric interviews have traditionally been used to diagnose depression and anxiety; however, carrying out psychiatric interviews during screening can be impractical given the large number of patients and limited resources. As such, self-report questionnaires can be useful alternatives.

One such questionnaire is the Hospital Anxiety and Depression Scale (HADS); however, reports of its effectiveness in screening these disorders have been mixed. Ibbotson et al. (1994) examined the performance of the HADS Unitary Scale against DSM-IV criteria in a mixed sample of cancer patients mixed both in terms of disease status and type of cancer. The measure was effective in correctly identifying affective disorders in those who were cancer free, in those who were currently receiving treatment and whose cancer was judged to be stable, although not in patients with progressive disease. Their study, however, did not examine the performance of the HADS subscales. In a palliative care setting Le Fevre et al. (1999) found that HADS was effective in identifying depression. In early-stage breast cancer, Hall et al. (1999) reported that the HADS anxiety and depression subscales were not effective in identifying patients who were clinically depressed and anxious. Their study, though, did not examine the performance of the HADS Unitary Scale. In a sample of cancer patients at different stages of illness Razavi et al. (1990) examined the performance of the HADS Unitary Scale and its subscales in identifying patients who were clinically depressed and anxious. Both subscales, especially the anxiety subscale, were less effective in correctly identifying such patients than the unitary scale.

These mixed findings are partly due to different cut-off scores used. The original authors recommended using a threshold of 11 on the HADS to include all probable cases of anxiety disorder and depressive illness, and a threshold of 8-10 to identify all possible cases of the 2 disorders. Hall et al. (1999), using a threshold of ≥ 11 , reported that the sensitivities of the HADS anxiety and depression subscales were 24.2% and 14.1%, respectively; lowering this threshold to ≥ 7 improved the sensitivity of the anxiety scale (72%), but not the depression scale (37.4%). Razavi et al. (1990) used various cut-off points on the subscales and reported lower sensitivity and specificity than the unitary scale. More specifically, a cut-off score of 19 on the full scale had 70% sensitivity and 75% specificity in identifying major depressive disorders, and a cut-off score of 13 had 75% sensitivity and 75% specificity in identifying both adjustment and major depressive disorders. On the other hand, using a cut-off score of 8 and 11 the HADS anxiety subscale had low sensitivity (64% and 54%, re-

spectively) and specificity (28% and 25%, respectively) in screening for adjustment disorders and major depressive disorders. Similarly, using a cut-off score of 7 and 9 the HADS depression subscale had low sensitivity (59% and 71%, respectively) and specificity (22% and 24%, respectively) in screening for adjustment disorders and major depressive disorders. Ibbotson et al. (1994) identified a cut-off point of 14 on the HADS Unitary Scale as the optimal threshold value for detecting affective disorders. Le Fevre et al. (1999) reported that the optimal threshold on the HADS Unitary Scale was 20.

Another alternative screening measure is the General Health Questionnaire (GHQ). Ibbotson et al. (1994) reported that GHQ-28 was effective in correctly identifying affective disorders in patients not being treated, with a cut-off score of > 7 ; however, it failed to do so in the overall sample, and among those who were disease free, those who had stable disease or progressive disease, and those who were in treatment. It has been proposed that the shorter version of this measure, GHQ-12, may be a better alternative to HADS (Hall et al. 1999); however, Hall et al. did not attempt to compare the performance of GHQ-12 to DSM-based clinical interviews. Le Fevre et al. (1999), on the other hand, showed that GHQ-12 was not effective in identifying psychiatric morbidity. Therefore, given these mixed findings the present study aimed to compare the accuracy of GHQ-12, and the HADS Unitary Scale and its subscales in identifying patients with affective disorders, including DSM major depression and generalized anxiety disorder.

METHOD

Sample

This study was part of a 2-phase study in breast cancer patients examining the effect of childhood adverse experiences on emotional adjustment following breast cancer surgery and at the 1-year follow-up. The study was funded by the Cancer Research Campaign, UK and carried out in 2 hospitals in the same region of England. The findings pertaining to the time of surgery have been published elsewhere (Clark et al. 2005; Clark et al. 2009; Salmon et al. 2006, 2007). The sample at the time of surgery consisted of 355 female patients with a diagnosis of primary breast cancer that had undergone mastectomy or wide local excision.

The present study reports the 1-year follow-up findings. In the present study, out of 355 female patients, 59 (17%) were excluded from the study because of metastases or recurrences, or they declined to participate. As such, 296

female breast cancer patients (83%) provided data relevant to this study. Mean age of the patients was 58 years (range: 25-83 years). 193 (65%) patients were married or living with a partner, 41 (14%) were divorced or separated, 40 (14%) were widowed, and 22 (7%) were single. In all, 130 (44%) patients were employed, 112 (38%) were retired, 23 (8%) were not working due to health problems, 10 (3%) were unemployed, 19 (6%) were homemakers, and 2 (1%) were students. The diagnosis was based on symptoms in 133 (45%) patients and screening results in 163 (55%). 188 (64%) patients underwent wide local excision and 108 (36%) had mastectomy. Two hundred and thirty (78%) patients received endocrine treatment, including tamoxifen or similar medication, 23 (8%) patients had chemotherapy, 123 (42%) had radiotherapy, 66 (22%) had both chemotherapy and radiotherapy, and 84 (28%) had no such treatment.

Procedure

Patients were contacted via telephone and invited to take part in the study. Those that agreed to participate were seen by the researcher at their home to administer the questionnaires and carry out the clinical interview.

In order to examine the accuracy of GHQ-12 (Goldberg and Hillier 1979) and HADS (Zigmond and Snaith 1983) in identifying patients with affective disorders, a standardized semi-structured psychiatric interview, the Schedule for Affective Disorders and Schizophrenia (SADS) (Endicott and Spitzer 1978) was used. This was supplemented by questions to generate a DSM-IV diagnosis. The interviewers were kept blind to the questionnaire data. All interviews were audio recorded. Interviewers were trained healthcare professionals, including psychologists and nurses. They were initially trained by audiotape demonstrations and practice. Then, via regular supervision from a psychiatrist and/or a psychiatric nurse, they were given feedback on the basis of their tape recordings and ratings.

More specifically, SADS was used to decide if a patient had DSM major depression or generalized anxiety disorder. Major depression was diagnosed if a participant reported a depressed mood for at least 2 weeks, accompanied by impairment in usual activities that was perceived by oneself and/or others and accompanied by at least 4 of the following symptoms: changes in weight or appetite, problems with sleep, loss of energy, loss of interest in usual activities, problems with sleep, loss of energy, loss of interest in usual activities, problems with concentration, thoughts of death or suicide, motor agi-

tation or retardation, retardation in speech or thought processes, and feelings of guilt, inadequacy, or disappointment. The use of SADS as a diagnostic tool makes it possible to diagnose generalized anxiety disorder on the basis of DSM criteria and/or Research Diagnostic Criteria (RDC) (Feighner et al. 1972). These 2 diagnostic criteria differ in terms of the time frame required to make a diagnosis of generalized anxiety disorder. That is, DSM-IV requires that the symptoms persist for 6 months, whereas RDC requires that they persist for 2 weeks. Nonetheless, evidence suggests that there is good agreement across the 2 criteria (Spitzer et al. 1975; Williams and Spitzer 1982). In the present study, we used the RDC to make a diagnosis of generalized anxiety disorder. Specifically, generalized anxiety disorder was diagnosed if a participant experienced generalized persistent anxiety for at least 2 weeks, accompanied by impairment in usual activities that was perceived by oneself and/or others and by symptoms from the following categories: feelings of restlessness, fatigue, problems with concentration, apprehensive expectation, irritability, muscular tension, and sleep problems.

HADS (Zigmond and Snaith 1983) consists of 7 items that measure the severity of depression and 7 items that measure the severity of anxiety. For each, patients are asked to choose 1 of 4 alternatives to indicate how they feel. This yields separate scores for anxiety and depression, as well as a total score. It is a widely used measure of emotional adjustment in patients with a physical illness. It has adequate internal consistency and construct validity (Zigmond and Snaith 1983). GHQ-12 (Goldberg and Hillier 1979) was developed for use in community samples to detect clinically significant emotional distress and has 12 items. For each item patients are asked to choose 1 of 4 alternatives to indicate how they feel. The answers were scored using the GHQ method (0,0,1,1). The original authors suggested that a score ≥ 2 indicates clinically significant distress (Goldberg 1972); however, others have used a threshold of 4 in breast cancer patients (Nosarti et al. 2002). In studies conducted with community samples this measure was reported to have good reliability and acceptable validity (Bennett 2000).

Statistical analysis

In order to examine whether the sample of the present study was biased a number of comparisons were made by computing t-tests. For these tests the level of significance was $P = 0.01$ instead of 0.05 in order to reduce the probability of making type 1 errors (i.e. rejecting the null hypothesis when it is true.)

The accuracy of GHQ-12 and HADS in identifying patients with affective disorders, including DSM major depression and generalized anxiety disorder, was examined by checking their sensitivity, specificity, positive predictive value, and misclassification rate.

We tested the accuracy of HADS and GHQ-12 in identifying affective disorders using the receiver operating characteristics (ROC) curve (Fawcett 2006, MEDCALC 2008). This is a graphical technique that plots the sensitivity of a test on the y-axis against specificity on the x-axis (Akobeng 2007). The area under the ROC curve (AUC) is a summary measure of the ROC curve and functions as a measure of the overall performance of a scale in relation to its accuracy in detecting affective disorders (Akobeng 2007; Walters 2005). If the AUC was < 0.70 , the accuracy of the scale in detecting affective disorders was unacceptable; if the AUC was between 0.70 and 0.80, the accuracy of the scale was considered to be fair; and if the AUC was between 0.80 and 0.90 or between 0.90 and 1.00, the accuracy of the scale was considered to be good or excellent, respectively (MEDCALC 2008). In addition to this summary measure, we also examined the accuracy of HADS and GHQ-12 in identifying affective disorders by paying attention to specific dimensions of accuracy, including sensitivity, specificity, positive predictive value, and misclassification rate.

The sensitivity of a scale pertains to the proportion of true cases of affective disorders in the sample (MEDCALC 2008). The higher the sensitivity, the lower the number of false negatives (those incorrectly identified as non-cases). The specificity of a scale pertains to the proportion of true non-cases in the sample (MEDCALC 2008). The higher the specificity, the lower the number of false positives (those incorrectly identified as true cases).

The positive predictive value (PPV) of a scale pertains to the probability of a score at or above a particular threshold being correctly identified as a true psychiatric morbidity (CLINLAB 2008). It is calculated by dividing the number of cases correctly identified (true positives) by the sum of true positives and false positives.

The misclassification rate (MR) pertains to the proportion of participants incorrectly classified by a scale as either false negative or false positive (CLINLAB 2008). MR is calculated by dividing the number of false negatives plus false positives by the total number in the sample.

All of these indices were calculated to examine the accuracy of the HADS Unitary Scale and its subscales, and

TABLE 1. The prevalence of clinically significant anxiety and depression across different measures.

Measure	Threshold	Total	Affective Disorders/ Emotional Distress (%)	Anxiety Frequency (%)	Depression Frequency (%)
SADS		260	28 (11%)	16 (6%)	22 (8%)
HADS Unitary Scale	≥19	296	41 (14%)		
HADS Unitary Scale	≥14	296	70 (24%)		
HADS Unitary Scale	≥13	296	76 (26%)		
HADS Unitary Scale	≥12	296	85 (29%)		
HADS Anxiety	≥11	296		45 (15%)	
HADS Depression	≥11	296			16 (5%)
HADS Anxiety	≥8	296		82 (28%)	
HADS Depression	≥8	296			44 (15%)
HADS Anxiety	≥7	296		99 (33%)	
HADS Depression	≥7	296			62 (21%)
GHQ-12	≥4	296	74 (25%)		
GHQ-12	≥3	296	75 (25%)		
GHQ-12	≥2	296	111(38%)		

GHQ-12 in detecting affective disorders as assessed by SADS. More specifically, the following comparisons were undertaken: 1) the performance of the HADS Unitary Scale and GHQ-12 was compared to affective disorders. Affective disorders consisted of the presence of either generalized anxiety disorder or DSM major depression or both; 2) the performance of GHQ-12 was compared to DSM major depression and generalized anxiety disorder; and 3) the performance of the HADS depression and anxiety subscales was compared to DSM major depression and generalized anxiety disorder, respectively.

RESULTS

Out of 296 patients who provided data relevant to this study, all completed HADS and GHQ-12, and 260 (88%) completed SADS. Comparisons across different measures included the data available for all relevant measures obtained from 255 patients (72%).

In order to examine whether the study sample was biased a number of comparisons were made. Those who completed and those who did not complete HADS and GHQ-12 in the present study did not differ in terms of their scores on the same measures at the time of surgery at $p \leq 0.01$ (GHQ-12: $T = 0.582$ [343], $P = 0.561$; HADS anxiety subscale: $T = 1.346$ [343], $P = 0.179$; HADS depression subscale: $T = 0.756$ [343], $P = 0.450$).

Moreover, those who completed and those who did not complete SADS in the present study did not differ in terms of HADS and GHQ-12 scores at the time of surgery at $p \leq 0.01$ (GHQ-12: $T = 0.877$ [343], $P = 0.381$; HADS anxiety subscale: $T = 1.241$ [343], $P = 0.215$; HADS depression subscale: $T = 0.928$ [343], $P = 0.354$). Similarly, those who completed and those who did not complete SADS in the present study did not differ in terms of their HADS and GHQ-12 scores at $p \leq 0.01$ (GHQ-12: $T = 1.509$ [294], $P = 0.132$; HADS anxiety subscale: $T = 1.773$ [294], $P = 0.077$; HADS depression subscale: $T = 2.439$ [294], $P = 0.015$).

Prevalence

The prevalence of clinically significant anxiety and depression across the different measures is given in Table 1. Overall, 28 patients (11%) were identified as having DSM major depression or generalized anxiety disorder based on SADS score. More specifically, 16 patients (6%) were identified as suffering from generalized anxiety disorder and 22 patients (8%) from DSM major depression. Based on the HADS subscales, depending on the threshold, the level of anxiety and depression ranged from 45 (15%) to 99 (33%) for anxiety and from 16 (5%) to 62 (21%) for depression. The same was true for the HADS Unitary Scale. Depending on the threshold the level of emotional distress ranged from 41 (14%) to

TABLE 2. Screening Performance of HADS Anxiety Subscale as compared to Generalized Anxiety Disorder (GAD) as assessed by SADS, n=255.

Area Under the Curve	Confidence Intervals	Threshold	Sensitivity %	GAD	False -ves	Specificity %	False +ves	•PPV %	••MR %
0.82	(0.68-0.95)	≥11	73.3	15	4	89.6	25	31	11
0.86	(0.77-0.94)	≥8	93.3	15	1	77.9	53	21	21
0.83	(0.74-0.92)	≥7	93.3	15	1	72.5	66	18	26

• PPV refers to Positive Predictive Value

•• MR refers to Misclassification Rate

85 (29%). The level of emotional distress, as measured by the GHQ-12, ranged from 74 (25%) to 111 (38%), depending on the threshold.

HADS subscales compared with clinical interviews

The accuracy of the HADS subscales in identifying patients with DSM major depression and generalized anxiety disorder was examined by checking their sensitivity, specificity, misclassification rate, and positive predictive value. In doing so, all the recommended cut-off points were examined (Tables 2 and 3). All of the HADS anxiety subscale thresholds were associated with ROC areas under the curve generally considered “good”, with correspondingly adequate sensitivities and specificities; however, the positive predictive values were much lower, reflecting the high number of false positives compared with true positives. The findings were similar for the HADS depression subscale in relation to DSM major depression, although 2 of the PPVs were somewhat higher than those for anxiety. One of the widely used thresholds (≥ 11) had an AUC in the unacceptable range.

HADS Unitary Scale compared with clinical interviews

The accuracy of the HADS Unitary Scale in identifying patients with affective disorders was examined by checking its sensitivity, specificity, misclassification rate, and positive predictive value. In doing so, all the recom-

mended cut-off points were examined (Table 4). Most of the HADS Unitary Scale thresholds were associated with ROC areas under the curve generally considered “good”, with correspondingly adequate sensitivities and specificities, however, the positive predictive values for these thresholds were much lower. One of the used thresholds (≥ 19) had low sensitivity, but much higher specificity and PPV.

GHQ-12 compared with clinical interviews

The accuracy of GHQ-12 in identifying patients with affective disorders, DSM major depression, and generalized anxiety disorder was examined by checking its sensitivity, specificity, misclassification rate, and positive predictive value. In doing so, the recommended or used cut-off points of 2 and 4 were examined (Goldberg 1972, Nosarti et al. 2002), as well as a cut-off point of 3.

Affective disorders: All of the GHQ-12 thresholds were associated with ROC areas under the curve generally considered “fair”, with correspondingly adequate sensitivities and specificities; however, the positive predictive values were lower (Table 4).

Generalized anxiety disorder: All of the thresholds were associated with ROC areas under the curve generally considered “fair”. One widely used threshold (≥ 2) had a higher sensitivity than the other thresholds, but had lower specificity. Moreover, the positive predictive values were low (Table 5).

TABLE 3. Screening Performance of HADS Depression Subscale as compared to DSM Major Depression as assessed by SADS, n = 255.

Area Under the Curve	Confidence Intervals	Threshold	Sensitivity %	DSM Major Dep.	False -ves	Specificity %	False +ves	•PPV %	••MR %
0.63	(0.49-0.77)	≥11	27.3	22	16	97.9	5	55	8
0.80	(0.68-0.92)	≥8	68.2	22	7	92.3	18	45	10
0.82	(0.72-0.93)	≥7	77.3	22	5	87.1	30	36	14

• PPV refers to Positive Predictive Value

•• MR refers to Misclassification Rate

TABLE 4. Screening Performance of the HADS Unitary Scale and the GHQ-12 , as compared to SADS affective disorders, n = 255. Thresholds ranging from ≥ 12 to ≥ 19 pertain to the HADS Unitary Scale and those ranging from ≥ 2 to ≥ 4 pertain to GHQ-12.

Area Under the Curve	Confidence Intervals	Threshold	Sensitivity %	DSM Affect. Disorder	False -ves	Specificity %	False +ves	•PPV %	••MR %
0.81	(0.70-0.92)	≥ 19	66.7	27	9	94.7	12	60	8
0.84	(0.75-0.92)	≥ 14	81.5	27	5	85.5	33	40	15
0.82	(0.74-0.91)	≥ 13	81.5	27	5	83.3	38	37	17
0.85	(0.77-0.92)	≥ 12	88.9	27	3	80.7	44	35	18
0.79	(0.69-0.89)	≥ 4	74.1	27	7	82.9	39	34	18
0.78	(0.68-0.88)	≥ 3	74.1	27	7	82.5	40	33	18
0.74	(0.64-0.84)	≥ 2	77.8	27	6	70.2	68	24	29

• PPV refers to Positive Predictive Value

•• MR refers to Misclassification Rate

DSM major depression: All of the thresholds were associated with ROC areas under the curve generally considered “fair”. One widely used threshold (≥ 4) had the same sensitivity as the other thresholds, but had much higher specificity. Moreover, the positive predictive values were low (Table 6).

DISCUSSION

The rate of affective disorders in general (11%), and DSM major depression (8%) and generalized anxiety disorder (6%) in particular identified by SADS was low. The levels of affective disorders identified by HADS and GHQ-12 varied, depending on the threshold used, but were generally higher than the levels estimated by SADS. In particular, using the HADS subscales and the cut-off scores of 8 on the HADS anxiety subscale and 7 on the HADS depression subscale, the levels of anxiety and depression were estimated to be 28% and 21%, respectively. These levels were similar to those reported by comparable studies in breast cancer that used self-report measures, including HADS and the Rotterdam Symptom Checklist. These studies showed that the levels of anxiety and depression were 21%-27% and 14%-21%, respectively. More specifically, in Nosarti et al.’s study (2002) the level of anxiety was estimated to be 21% using

a cut-off score of 8 on the HADS anxiety subscale and that of depression was 14% using a cut-off score of 7 on the HADS depression subscale.

The present study examined the accuracy of HADS and GHQ-12 in detecting affective disorders, including DSM major depression and generalized anxiety disorder as measured by SADS. During the process of deciding on the optimal thresholds, considering the trade-off between sensitivity and specificity, we paid more attention to sensitivity. This was because we thought it was important to keep the number of false negatives as low as possible during the screening process. Otherwise, during this process healthcare professionals may miss patients who are emotionally distressed and need psychological help.

When the performance of the HADS subscales was compared to SADS, the findings suggested that a cut-off score ≥ 8 on the HADS anxiety subscale with 93.3% sensitivity and 77.9% specificity was the optimal threshold for detecting generalized anxiety disorder. In relation to the detection of DSM major depression, the optimal threshold on the HADS depression subscale was ≥ 7 , with 77.3% sensitivity and 87.1% specificity. Overall, these findings suggest that the HADS subscales are able to accurately detect DSM major depression and generalized anxiety disorder. These findings are somewhat inconsistent

TABLE 5. Screening Performance of GHQ-12 as compared to Generalized Anxiety Disorder (GAD) as assessed by SADS n = 255.

Area Under the Curve	Confidence Intervals	Threshold	Sensitivity %	GAD	False -ves	Specificity %	False +ves	•PPV %	••MR %
0.73	(0.59-0.87)	≥ 4	66.7	15	5	79.6	49	17	21
0.73	(0.59-0.87)	≥ 3	66.7	15	5	79.2	50	17	22
0.70	(0.57-0.84)	≥ 2	73.3	15	4	67.5	78	12	32

• PPV refers to Positive Predictive Value

•• MR refers to Misclassification Rate

TABLE 6. Screening Performance of the GHQ-12 as compared to DSM Major Depression as assessed by the SADS, N = 255.

Area Under the Curve	Confidence Intervals	Threshold	Sensitivity %	DSM Major Dep.	False -ves	Specificity %	False +ves	•PPV %	••MR %
0.80	(0.69-0.90)	≥4	77.3	22	5	82.0	42	29	18
0.79	(0.69-0.90)	≥3	77.3	22	5	81.5	43	28	19
0.73	(0.62-0.84)	≥2	77.3	22	5	69.1	72	19	30

• PPV refers to Positive Predictive Value

•• MR refers to Misclassification Rate

ent with those of Hall's (1999), whose study was comparable to ours in terms of sample characteristics, but not in terms of time frames. The present study targeted affective disorders at the end of the first post-surgery year, whereas Hall (1999) targeted those around the time of surgery. Hall et al. (1999), using the threshold of ≥ 7 , reported the HADS anxiety subscale had 72% sensitivity, but that of the HADS depression subscale was too low (37.4%); that is, their study found that the threshold of 7 was reasonably accurate in screening for anxiety, but not depression.

When the performance of the HADS Unitary Scale was compared to SADS in detecting overall affective disorders, the findings suggested that a cut-off score ≥ 12 was the optimal threshold. This had 88.9% sensitivity and 80.7% specificity. Overall, this finding and the finding that a score ≥ 8 and ≥ 7 are the optimal thresholds on the HADS anxiety and depression subscales in detecting generalized anxiety disorder and DSM major depression, respectively, are inconsistent with the general finding reported by Razavi et al. (1990) that the HADS Unitary Scale is more effective in accurately detecting generalized anxiety disorder and major depression.

Our findings are also not consistent with Ibbotson et al.'s study (1994) in relation to the exact threshold on the HADS Unitary Scale. Our study suggests that a score ≥ 12 on the HADS Unitary Scale should be the optimal threshold in detecting affective disorders, whereas Ibbotson et al. (1994) suggest that the threshold of 14, with 80% sensitivity, should be considered the optimal threshold.

When the performance of GHQ-12 was compared to SADS the areas under the curve were within the fair range and were lower than those of the HADS Unitary Scale and its subscales. Nevertheless, the optimal threshold was ≥ 2 , with 77.8% sensitivity and 70.2% specificity. In detecting generalized anxiety disorder the optimal threshold was ≥ 2 , with 73.3% sensitivity and 67.5

specificity. In detecting DSM major depression the optimal threshold was ≥ 4 , with 77.3% sensitivity and 82% specificity. These findings suggest that GHQ-12 can be used as a screening measure to accurately detect affective disorders, DSM major depression, and generalized anxiety disorder, and that a cut-off score ≥ 2 on GHQ-12 can be used as the optimal threshold for detecting affective disorders and generalized anxiety disorder, and a cut-off score ≥ 4 can be used as the optimal threshold for detecting DSM major depression. The threshold of ≥ 2 is consistent with the suggestions of the original authors that a score ≥ 2 and above indicates clinically significant distress (Goldberg 1972) and a threshold ≥ 4 is consistent with its use in breast cancer patients (Nosarti et al. 2002).

The performance of GHQ-12 in detecting affective disorders confirms Hall et al.'s (1990) recommendation to use GHQ-12 as a screening measure. This is also consistent with the findings of another study (Ibbotson et al. 1994) that reported GHQ-28 was effective in correctly identifying affective disorders in patients that were off treatment with a cut-off score > 7 ; however, it failed to do so in the overall sample, and among those who were disease free, those who had stable disease or progressive disease, and those who were in treatment. Similarly, in the present study we obtained the same findings in a sample that excluded patients with recurrences and metastases, in which most of the patients completed their treatment.

Variation in the performance of GHQ-12 across different types of affective disorders, including DSM major depression and generalized anxiety disorder, is worth noting. That is, the sensitivity and specificity of GHQ12 in detecting DSM major depression were higher than in detecting generalized anxiety disorder. The original authors of GHQ-12 claimed that this scale measured emotional distress. An analysis of the GHQ-12 items indicates that the scale focuses on general symptoms of psychiatric morbidity, including anxiety and depression,

but more so on depression; therefore, it is not surprising that it can detect major depression more effectively than generalized anxiety disorder.

When the performance of GHQ-12 and HADS is compared our findings suggest that the former detects affective disorders in early-stage breast cancer patients less accurately than the latter. More specifically, the areas under the curve for HADS were generally within the good range, whereas for GHQ-12 they were generally within the fair range, and the optimal threshold on GHQ-12 (≥ 2) had 77.8% sensitivity, whereas the optimal threshold for the HADS Unitary Scale (≥ 12) had 88.9% sensitivity. This is consistent with the findings of Le Fevre et al. (1999), who also reported that GHQ-12 was less effective than HADS in identifying psychiatric morbidity.

In the present study GHQ-12 may have been less effective in detecting affective disorders because of the different time frames used across different measures of affect (Hall et al. 1999). The GHQ-12 items measure mood during the past few weeks. HADS, on the other hand, measures mood during the past week. SADS measures the last 2 weeks, including the present. Therefore, one would expect better concordance between HADS and SADS than between GHQ-12 and SADS. That is, GHQ-12 could have identified a larger number of patients who experienced emotional problems (111 patients [38%] at ≥ 2) than HADS (85 patients [29%] at ≥ 12), but whose symptoms disappeared or began to disappear within the last 2 weeks, as targeted by SADS. A similar picture has been observed in other studies in relation to the Present State Examination and HADS (Hall et al. 1999).

In general, it is noteworthy that optimal thresholds on HADS and GHQ-12 identified in the present study

have relatively high sensitivity. This is encouraging because it is important to keep the number of false negatives as low as possible; however, what is equally important is the positive predictive values. For the optimal thresholds, these unfortunately were low in the present study, ranging from 12% to 36%. In other words, using these thresholds the probability of an identified case being a true case is relatively low. This means that there is a good chance that a higher number of patients will be identified as depressed or anxious than actually exists, and that this will likely increase the burden on the clinical staff. This may be due to the small number of patients identified as depressed and anxious by SADS.

Moreover, the present study included breast cancer patients at 1-year post surgery (including mastectomy or wide local excision) in 2 hospitals in 1 region of England. This study excluded patients with metastases or recurrences. Cultural and clinical differences are likely to reduce the generalizability of our results. More specifically, these findings may not be generalizable to patients with metastases or recurrences, or to those who are receiving treatment for breast cancer in other parts of the UK and the world. Similarly, these findings will not be generalizable to other cultures.

Overall, keeping these limitations in mind, the HADS Unitary Scale and its subscales can be used in clinical settings as a screening instrument for detecting affective disorders in breast cancer patients. The threshold for the HADS Unitary Scale should be ≥ 12 and should be ≥ 7 and 8 for the depression and anxiety subscales, respectively. GHQ-12 can also be used as a screening instrument to detect affective disorders, generalized anxiety disorder, and DSM major depression. The threshold should be ≥ 2 for affective disorders and generalized anxiety disorder, and ≥ 4 for DSM major depression.

REFERENCES

- Akobeng A K (2007) Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatr*, 96: 644-647.
- Bennett P (2000) *Introduction to Clinical Health Psychology*. Open University Press: England.
- Burgess C, Cornelius V, Love S et al. (2005) Depression and anxiety in women with early breast cancer: five year observational cohort study. *BMJ*, 330: 702-705.
- Clark L, Holcombe C, Hill J, et al. (2005). What do patients bring to the consultation? Professional and social support in breast cancer: the role of childhood abuse. *Breast Cancer Res Treatment*, 94: S276-S276.
- Clark L, Holcombe C, Hill J, et al. (2009) The perception of support received from breast care nurses by depressed patients following a diagnosis of breast cancer. *Ann R Coll Surg Engl* 91: 43-45.
- Clinlab. Parameters for studies of diagnostic accuracy. http://clinlab.uch.edu/PathlinkData/Evidence_Based_Medicine/Dx_Parameters.htm 15.09. 2008.
- Endicott J, Spitzer R L (1978) A diagnostic interview: the Schedule for Affective Disorders and Schizophrenia. *Arch Gen Psychiatry*, 35: 837-844.
- Fallowfield L, Hall A, Maguire P, Baum M (1990) Psychological outcomes of different treatment policies in women with early breast cancer outside a clinical trial. *BMJ*, 301: 575-580.
- Fawcett T (2006) An introduction to ROC Analysis. *Patterns Letters Recognition*, 27: 861-874.
- Feighner JP, Robins E, Guze SB (1972) Diagnostic criteria for use in psychiatric research. *Arch Gen Psychiatr*, 26: 57-63.
- Goldberg DP (1972) *The detection of psychiatric illness by questionnaire*. London: Oxford University Press.

- Goldberg DP, Hillier VF (1979) A scale version of the General Health Questionnaire. *Psychol Med*, 9: 139-145.
- Goldberg JA, Scott RN, Davidson PM, et al. (1992) Psychological morbidity in the first year after breast surgery. *Eur J Surg Oncol*, 18: 327-331.
- Hall A, Hern A, Fallowfield L (1999) Are we using appropriate self-report questionnaires for detecting anxiety and depression in women with early breast cancer? *EJC*, 35: 79-85.
- Ibbotson T, Maguire P, Selby P, et al. (1994) Screening for anxiety and depression in cancer patients: the effects of disease and treatment. *EJC*, 30A: 37-40.
- Le Fevre P, Devereux J, Smith S, et al. (1999) Screening for psychiatric illness in the palliative care inpatient setting: a comparison between the Hospital Anxiety and Depression Scale and the General Health Questionnaire-12. *Palliat Med*, 13: 399-407.
- Maguire P (1985) Barriers to psychological care of the dying. *Br Med J*, 291: 1711-1713.
- Medcalc, ROC curve analysis: introduction. <http://www.medcalc.be/manual/roc.php> 10.01.2008.
- Nosarti C, Roberts J V, Crayford T, et al. (2002) Early psychological adjustment in breast cancer patients: A prospective study. *J Psychosom Res*, 53: 1123-1130.
- Razavi D, Delvaux N, Farvacques C, et al. (1990) Screening for adjustment disorders and major depressive disorders in cancer in-patients. *Br J Psychiatry*, 156: 79-83.
- Salmon P (2000) *Psychology of medicine and surgery: a guide for psychologists, counsellors, nurses and doctors*. John Wiley and Sons Ltd: UK.
- Salmon P, Manzi F, Valori RM. (1996) Measuring the meaning of life for patients with incurable cancer: the Life Evaluation Questionnaire (LEQ). *EJC*, 32A: 755-760.
- Salmon P, Hill J, Krespi R, et al. (2006) The role of child abuse and age in vulnerability to emotional problems after surgery for breast cancer. *EJC*, 42: 2517-2523.
- Salmon P, Holcombe C, Clark L, et al. (2007) Relationships with clinical staff after a diagnosis of breast cancer are associated with patients' experience of care and abuse in childhood. *J Psychosom Res*, 63: 255-262.
- Spitzer R L, Endicott J, Robins E, et al. (1975) Preliminary Report of the Reliability of Research Diagnostic Criteria (RDC) Applied to Psychiatric case Records. Ravens Press.
- Walter SD (2005) The partial area under the summary ROC curve. *Stat Med*, 24: 2025-2040.
- White CA, Macleod U (2002) ABC of psychological medicine: cancer. *BMJ*, 325: 377-382.
- Williams JBW, Spitzer RL (1982) Research diagnostic criteria and DSM-III. *Archives of General Psychiatry*, 39: 1283-1289.
- Zigmond AS, Snaith RP (1983) Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*, 67: 361-370.