INTERNATIONAL EXPERIENCES WITH THE HOSPITAL ANXIETY AND DEPRESSION SCALE—A REVIEW OF VALIDATION DATA AND CLINICAL RESULTS

CHRISTOPH HERRMANN

(Received 28 May 1996; accepted 26 June 1996)

Abstract—More than 200 published studies from most medical settings worldwide have reported experiences with the Hospital Anxiety and Depression Scale (HADS) which was specifically developed by Zigmond and Snaith for use with physically ill patients. Although introduced in 1983, there is still no comprehensive documentation of its psychometric properties. The present review summarizes available data on reliability and validity and gives an overview of clinical studies conducted with this instrument and their most important findings. The HADS gives clinically meaningful results as a psychological screening tool, in clinical group comparisons and in correlational studies with several aspects of disease and quality of life. It is sensitive to changes both during the course of diseases and in response to psychotherapeutic and psychopharmacological intervention. Finally, HADS scores predict psychosocial and possibly also physical outcome. Copyright © 1997 Elsevier Science Inc.

Keywords: Anxiety; Depression; Hospital Anxiety and Depression Scale; Medical patients; Psychological screening; Review article.

INTRODUCTION

The Hospital Anxiety and Depression Scale (HADS) was developed by Zigmond and Snaith [1] in 1983 to provide clinicians and scientists with a reliable, valid, and practical tool for identifying and quantifying the two most common forms of psychological disturbances in medical patients. It was designed with special attention to some specific issues especially relevant for the setting of somatic medicine. The scale was limited to 14 items, which makes it easy to administer and well accepted. Severely psychopathological symptoms are not covered. This is thought to improve acceptability and make the scale more sensitive to mild forms of psychiatric disorders, thus avoiding the "floor effect," which is frequently observed when psychiatric questionnaires are used with medical patients. Furthermore, the investigators omitted physical indicators of psychological distress such as headache or weight loss, which could give false positive results if they were in fact due to an underlying medical illness. Finally, scale scores should not be overresponsive to transient fluctuations in state which may occur in situations such as coming to a clinic. On the other
hand, they should respond well to mood changes which may occur during the course of a disease. Thus, the scale should be a prolonged state rather than trait measure.

Because most other scales have considerable shortcomings with respect to at least one of these criteria, the scale is now widely used in England as well as in several other countries. It is, however, very surprising that most clinicians and investigators still rely mainly on Zigmond and Snaith's original study [1], although it provides only very sparse data on reliability and validity. There have now been more than 200 publications reporting original experiences with the instrument in approximately 35,000 persons, but still there is no comprehensive documentation of its psychometric properties. Even the English test manual [2] does not give sufficient information. The largest systematic validation study has been performed with the German HADS version in 6200 patients and control subjects but it is only available in German and cannot simply be generalized to other national versions. One purpose of this review, therefore, is to summarize available validation data for the English and German HADS versions which will be shown to be almost equivalent. In addition, this article gives an overview of the extensive clinical experiences with different national versions of the scale in most medical settings. The results of these studies might be considered a clinical validation of the scale. They also provide an international perspective on the current state of psychosomatic research.

Relevant literature was retrieved over a period of 8 years by means of repeated searches in Medline, CancerCD, and Current Contents (Social & Behavioral Sciences) as well as by personal correspondence with Dr. Snaith. In addition, reference lists of the identified publications were screened for additional studies which might report data about the HADS. Data from the German validation sample were reanalyzed using the original data file characterized in the German manual [3].

ACCEPTABILITY, RELIABILITY AND VALIDITY OF THE HADS

Acceptability

The scale is generally well accepted by patients and nonpatients alike. Many studies report response rates of 100% in well-motivated study patients. Under routine conditions, some patients will not be able or willing to complete the scale. Nevertheless, even in an acute cardiological care setting we found a high acceptance of 95% in a sample of 531 patients [4]. Similar results have been reported by other investigators [e.g. 5–9]. The scale can be completed in 2–6 minutes and—with practice—scored in 1 minute, which makes it easy for nonpsychiatric doctors or nurses to handle.

Reliability

All items are scored on a 4-point scale from 0 to 3. Item statistics for the German HADS version [3, 10] are displayed in Table 1. Similar results have been obtained by other investigators using different national HADS versions [1, 11–17]. With few exceptions, the results show satisfactory or good item-total correlations within the two subscales. Internal consistencies (Cronbach alphas) of the English and German versions are also acceptable at 0.80 to 0.93 for the anxiety and 0.81 to 0.90 for the depression subscales [3, 11, 15, 18]. Similar or slightly lower values have been observed in smaller studies with other HADS versions.
Table I.—Item characteristics and reliability of the German HADS version ($n = 6200$)

<table>
<thead>
<tr>
<th>Item characteristics</th>
<th>Anxiety subscale</th>
<th>Depression subscale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means</td>
<td>0.65–1.33</td>
<td>0.44–1.16</td>
</tr>
<tr>
<td>Corrected item-total correlations ($r = 5338$)</td>
<td>0.52–0.68</td>
<td>0.49–0.72</td>
</tr>
<tr>
<td>Internal consistency ($n = 5338$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cronbach’s $\alpha$</td>
<td>0.80</td>
<td>0.81</td>
</tr>
<tr>
<td>Retest reliability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–2 weeks ($n = 79$)</td>
<td>0.84</td>
<td>0.85</td>
</tr>
<tr>
<td>&gt;2–6 weeks ($n = 111$)</td>
<td>0.73</td>
<td>0.76</td>
</tr>
<tr>
<td>&gt;6 weeks ($n = 901$)</td>
<td>0.70</td>
<td>0.70</td>
</tr>
</tbody>
</table>

*Subjects with complete answers to all of the 14 HADS items.

[16, 17, 19]. Retest reliability shows a high correlation, $r > 0.80$, after up to 2 weeks, which decreases with longer time intervals (Table 1) [8, 20–24]. Thus, unlike typical state instruments, the HADS is stable enough to withstand situational influences. On the other hand, over a longer span of time it is less stable than typical trait scales and should respond to mood changes.

Validity

Factorial validity. Factor analysis of the English and German HADS versions resulted in nearly identical solutions with one depression and one anxiety factor. These factors remain stable across subgroups [3, 18], correlate highly with the corresponding subscales ($r > 0.90$) [3], and explain about 50% of variance (Table II). A two-factor model has also been reported for the Swedish version [19, 25], although a third factor emerged in relatively healthy patients [19]. Smaller studies with the Chinese and French versions [12, 15] have also found three-factor solutions. Nevertheless, in the Chinese study, the first two factors also clearly resemble those from England and Germany (Table II), whereas only two items load significantly on the third factor. Strikingly, all the three models displayed in Table II show unexpected loadings of the fourth anxiety item on the depression factor. The fifth depression item has relatively low loadings on the depression factor. Thus, despite the generally very satisfactory confirmation of the two subscales postulated, future revisions may have to reconsider these items.

After orthogonal (varimax) rotation factor correlations with the noncorresponding subscales are low ($< 0.4$) [3] indicating that the subscales can be conceptualized as two theoretically independent dimensions of mood. On the other hand, oblique rotation which takes into account the real coexistence of anxious and depressed symptoms in most patient groups results in factor correlations of about 0.5. One could conclude that the HADS theoretically allows one to make separate assessments of anxiety and depression but that in reality there will be some overlap.

Discriminant and concurrent validity. There has been some discussion of whether the anxiety and depression subscales really measure different aspects of mood. Some investigators [12, 26] have preferred the HADS total score as
Table II.—HADS structure matrix after oblique rotation in the English, German, and Chinese HADS versions

<table>
<thead>
<tr>
<th>Anxiety items</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>English version [18] (n = 568)</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>German version (n = 5338)</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Chinese version [15] (n = 141)</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Depression items

| 2             | ++       | -        | ++                             | -        | +        | ++                         | -        |
| 4             | ++       | -        | ++                             | -        | ++       | -                          | -        |
| 6             | ++       | -        | ++                             | -        | ++       | -                          | -        |
| 8             | +        | -        | +                              | -        | +        | -                          | -        |
| 10            | -        | -        | +                              | -        | -        | -                          | -        |
| 12            | ++       | -        | ++                             | -        | ++       | -                          | -        |
| 14            | +        | -        | +                              | -        | +        | -                          | -        |

Explained variance

<table>
<thead>
<tr>
<th>English version [18] (n = 568)</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>53%</th>
</tr>
</thead>
<tbody>
<tr>
<td>German version (n = 5338)</td>
<td>Factor 1</td>
<td>Factor 2</td>
<td>48%</td>
</tr>
<tr>
<td>Chinese version [15] (n = 141)</td>
<td>Factor 1</td>
<td>Factor 2</td>
<td>47%*</td>
</tr>
</tbody>
</table>

Factor loadings: −, ≤0.50; +, >0.50. ++, >0.70.

* Another 9% of variance explained by a third factor with significant loadings of items 10 and 11.

a general measure of distress, others [27] have found no significant correlation between the two subscales. In fact, the mean correlation between anxiety and depression subscales from 18 separate studies with n = 8160 is r = 0.63 [3, 8, 12, 15, 20, 22, 27–38]. Thus, subscale scores are clearly correlated in most patient groups. Nevertheless, there is sufficient evidence that both subscales differ in a clinically meaningful way (see below). It has also been demonstrated by several studies that the HADS anxiety subscale shows significantly higher correlations with observer ratings and self-assessment questionnaires for anxiety (versus depression), while the depression subscale correlates better with external criteria for depression (versus anxiety) [1, 3, 4, 8, 12, 20, 27, 29, 31, 34, 39–48]. Partial correlation shows that the correlations of each subscale with the corresponding observer ratings are independent of the other subscale [45]. Finally, the external criteria for anxiety and depression are also substantially correlated with each other in most studies. Clearly, the correlation between HADS anxiety and depression subscales is mainly due to a real coincidence of anxious and depressed symptoms in the patient groups and only to a lesser extent to inadequacies of the instrument. In absolute values, correlations of both HADS subscales with their corresponding criteria are satisfactory or good [3].

Sensitivity and specificity for identifying psychiatric “cases.” One of the main purposes of the HADS is to identify psychological disturbances in medical patients. Like any other self-rating instrument it can indicate that a particular patient is probably a psychiatric “case” of anxiety or depression. It does, however, not allow one to make definite diagnoses and gives a dimensional rather than
categorical representation of mood. For identifying subjects likely to be anxious or depressed, the ordinal scale score [1, 49] has to be dichotomized at a previously defined cutoff point. The ability of a screening test to differentiate between subjects with versus those without a specific disorder at a given cutoff can be characterized by its sensitivity and specificity. In general, increasing the cutoff also increases specificity while reducing sensitivity. Thus, any choice of a cutoff is a compromise between sensitivity and specificity. If it is necessary to identify almost all “cases,” even at the cost of a relevant number of false positives, it will be useful to choose a low cutoff. In other settings, where, for example, only few severely disordered patients can be offered an expensive intervention, a higher cutoff will be more appropriate.

There is no single, generally accepted cutoff score for the HADS. In their original study, Zigmond and Snaith [1] recommended two cutoff scores for both subscales: 7/8 for possible and 10/11 for probable anxiety or depression (with possible ranges of 0–21 for each subscale). In their manual [2], they also proposed a third cutoff of 14/15 for “severe” disorder, without presenting empirical data for its usefulness. Most investigators have used one of the lower two cutoffs, and others have recommended different scores or even introduced new cutoffs for the total HADS sum score. At the same time, several external criteria have been used as “gold standards.” It is therefore not possible to give unequivocal general values for sensitivity and specificity of the instrument. Seventeen studies for the English version report average sensitivities and specificities of 0.8 or higher [1, 27, 30, 42, 45, 50–61]. These results compare favorably with validity data for other screening tests (such as the exercise ECG in detecting coronary artery disease). Several studies with international HADS versions found comparable or only slightly worse results [3, 16, 17, 39, 41, 62–65]. However, unacceptably low specificities of less than 0.5 were observed in three studies with Asian or Australian patients [61, 66, 67]. These are possibly due to patient characteristics (e.g., poststroke patients) or weak external criteria, since other investigators found very satisfactory specificities with the same national HADS versions (i.e., Urdu, Arab, and English).

Although sensitivity and specificity of the HADS for detecting psychiatric “cases” are good, it has been shown that the scale performs poorly when it is used for making a specific diagnosis of major depression in medical or psychiatric patients [59]. The author of that study points out that the positive predictive value (i.e., the percentage of subjects with HADS scores above a cutoff who are clinically found to have major depression) is very low. This seems surprising, because sensitivity and specificity in that study are similar to those reported by others. However, one has to recall how screening tests work; that is, they always give a relatively high percentage of false positive results in populations with very low prevalences or prior probabilities of the criterion, whereas in populations with very high prevalences there are high percentages of false negative results. One typical example is the frequently observed finding of false negative exercise ECGs in young women with chest pain, in whom the prevalence of coronary disease is extremely low. The low positive predictive value of the HADS in Silverstone’s group of medical patients [59] can simply be explained by the same effect, as only 6% percent of the medical patients had major depression.
Table III.—Prevalences of abnormal HADS scores in different German patient and control groups

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Abnormal anxiety (HADS-A &gt;10)</th>
<th>Abnormal depression (HADS-D &gt;8)</th>
<th>Abnormal anxiety and/or depression</th>
<th>Prevalence ratio*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical students</td>
<td>136</td>
<td>9%</td>
<td>2%</td>
<td>10%</td>
<td>4.5</td>
</tr>
<tr>
<td>Healthy controls (general population)</td>
<td>152</td>
<td>7%</td>
<td>5%</td>
<td>10%</td>
<td>1.4</td>
</tr>
<tr>
<td>Nonpatients with health complaints</td>
<td>55</td>
<td>33%</td>
<td>13%</td>
<td>38%</td>
<td>2.5</td>
</tr>
<tr>
<td>Oncology outpatients</td>
<td>73</td>
<td>10%</td>
<td>21%</td>
<td>25%</td>
<td>0.5</td>
</tr>
<tr>
<td>Cardiological patients</td>
<td>5579</td>
<td>19%</td>
<td>17%</td>
<td>26%</td>
<td>1.1</td>
</tr>
<tr>
<td>Neurological outpatients</td>
<td>11</td>
<td>18%</td>
<td>27%</td>
<td>27%</td>
<td>0.7</td>
</tr>
<tr>
<td>Patients on hemodialysis</td>
<td>18</td>
<td>22%</td>
<td>50%</td>
<td>50%</td>
<td>0.4</td>
</tr>
<tr>
<td>Back pain patients</td>
<td>70</td>
<td>36%</td>
<td>29%</td>
<td>47%</td>
<td>1.2</td>
</tr>
<tr>
<td>Fibromyalgia patients</td>
<td>33</td>
<td>39%</td>
<td>39%</td>
<td>52%</td>
<td>1.0</td>
</tr>
<tr>
<td>Psychiatric patients</td>
<td>69</td>
<td>56%</td>
<td>48%</td>
<td>71%</td>
<td>1.2</td>
</tr>
</tbody>
</table>

* Prevalence of abnormal anxiety divided by prevalence of abnormal depression.

However, in that study, prediction was also poor in the psychiatric group. In these patients HADS obviously did not discriminate between major depression and depressed mood occurring as a symptom of other psychiatric disorders. One should therefore not use HADS for making a diagnosis of a specific type of depression. Nevertheless, the author of that study also recommends the use of the scale for its original purpose.

Several studies document that the HADS detects anxiety and depression significantly better than non-psychiatric physicians [e.g., 43, 51, 55].

Reference values. A different way of interpreting individual HADS scores in a more dimensional manner is to compare them with established reference values. Unfortunately, there are no documented standard values for the English version. There are, however, numerous studies reporting distributions of scores for a vast variety of patient groups which can be used as preliminary standards (see below). Percentile and t-standardized scores are given in the German HADS manual [3]. For cardiological patients, they have been adjusted for age and gender, whereas only gender-specific standard values are available for healthy persons. Because formal characteristics are almost identical for the English and German versions, and HADS scores have also been found identical in comparable groups of cardiological patients from England and Germany [4], the German standard values should also apply to English heart patients. For other patient groups it might make more sense to use the standard values for healthy persons as a reference.

Construct validity. One aspect of construct validation is to compare mean scores or prevalences of positive HADS scores in groups with known differences in clinically relevant anxiety or depression. For the German validation sample (n=6200) prevalences of abnormal anxiety (cutoff 11) and depression (cutoff 9) [3] are displayed in Table III. Abnormal anxiety is observed in 7% of healthy persons and 56% of psychiatric patients. Patients with physical disease have anxiety prevalences of 10–22% on the HADS, and those with chronic
pain as well as nonpatients with different complaints are significantly anxious at 33–39%. Prevalences of abnormal depression scores range from 2% in medical students to 48% in psychiatric patients and 50% in patients with terminal renal insufficiency. Interestingly, the ratio of anxious to depressed patients differs considerably among groups and ranges from 0.4 in patients with chronic disabling diseases to 2.5 in nonpatients with different complaints and 4.5 in medical students. At least one abnormal HADS score was observed between 10% of medical students and healthy controls and 71% of psychiatric patients. The latter figure is lower than the sensitivity of the scale because not all psychiatric patients suffered from depressive or anxiety disorders.

_Treatment validation._ One important aspect of psychological screening tests is their ability to reflect changes in the criterion variable which occur during psychosocial or psychopharmacological intervention. This “treatment validation” of the HADS, as requested by Meakin [56], has been accomplished by several treatment studies. These included psychiatric patients with major depression [42, 68] and neurotic disorders [69] who were treated with psychotropic drugs and/or psychotherapy as well as patients with cancer [70–72] or heart disease [73–80] who received psychosocial interventions or rehabilitation programs. In all of these studies, HADS scores were significantly reduced as a consequence of the interventions. Some of the studies also showed significant improvements in HADS scores when comparing treated patients with control groups [71–73, 75–77, 80]. In two studies, the best results were obtained in patients with high initial HADS scores [74, 77].

_Comparisons with other self-rating scales._ The ability of the HADS to detect psychiatric “cases” is similar to that reported for other scales, with different studies showing slight advantages of one or another instrument in different patient groups [e.g., 4, 51, 53, 56, 57, 81]. There are no indications from the literature that other self-rating scales differentiate significantly better between anxiety and depression. Formal characteristics of the HADS do not differ significantly from those reported for other instruments, with the exception of its favorable retest reliability (see above). In medical patients, HADS scores seem to be more normally distributed than for example those of the State–Trait Anxiety Inventory (STAI) [47] or the Beck Depression Inventory (BDI) [46, 82], which, in addition, exhibits a relevant floor effect. Furthermore, the BDI, like many other scales, contains several somatic indicators of depression, whereas none of the HADS items asks for such symptoms, making this scale unsusceptible to symptoms of physical disease [1]. In one study, HADS anxiety scores but not STAI (trait) or BDI successfully predicted normal coronary angiograms in uni- and multivariate analyses of chest pain patients [82].

_Preliminary conclusions_ The HADS is well accepted by patients. Formal characteristics are satisfactory or good. The two-dimensional structure has been confirmed by factor analyses, despite a relevant correlation between the anxiety and depression subscales. Sensitivity and specificity for detecting “cases” are acceptable, although high depression scores _per se_ do not allow one to make a diagnosis of major depression. The HADS discriminates well between samples with high, medium,
A. Cross-sectional studies
1. Validation and feasibility studies
2. Simple screening studies (prevalence, mean values)
3. Group comparisons:
   - Within diagnostic groups
   - Across diagnostic groups (e.g., functional versus somatic causes of symptoms)
   - Across disciplines
4. Correlations:
   - With severity and symptoms of physical illness
   - With personality variables or quality of life
   - With new self-rating questionnaires (HADS as a validity criterion)

B. Longitudinal studies
1. Natural course of anxiety and depression
2. Effects on HADS of:
   - Disease parameters
   - Medical or surgical interventions
   - Life events
   - Information and decision making
   - Psychosocial and psychopharmacological interventions
3. Prediction of outcome by baseline HADS scores

and low prevalences of anxiety or depressive disorders. Standard values are available for some patient groups. Anxiety and depression scores can be reduced significantly by psychosocial interventions or drug treatment.

CLINICAL EXPERIENCES WITH HADS

After its initial validation, the HADS has been used in routine care as well as in multiple types of clinical investigations (Tables IV and V) in and outside the United Kingdom (Table VI). The following sections show a breakdown of studies by study type and medical discipline, irrespective of the countries in which the studies were performed. Later we shall take a separate look at those studies which were performed with different national versions outside the United Kingdom.

Cross-sectional studies used the scale for simple screening purposes, for comparing mean levels or prevalences of anxiety and depression in different patient (sub)groups or for correlating scale scores with parameters of somatic disease, personality, or quality of life. Some more recent studies used the HADS as a validity criterion in evaluating new questionnaire instruments. In longitudinal studies, HADS scores have been used as simple descriptive parameters for the course of anxiety and depression in specified patient groups, as dependent (outcome) variables, and also sporadically as predictors of psychosocial or physical outcome.

Cross-sectional studies

Table V shows the types of data reported in the available publications by medical specialties in which these studies were conducted. As can be seen, the largest single group of studies consists of cross-sectional group comparisons (Table V, Category A3). These comparisons include within-diagnosis compari-
sons (e.g., men versus women, stable versus progressive disease, silent versus symptomatic myocardial ischemia, etc.), across-diagnose comparisons (e.g., Crohn’s disease versus functional abdominal complaints), and across-discipline comparisons (such as the one displayed in Table III) including comparisons with healthy control subjects. From these studies, as well as from some simple screening studies (Table V, Category A2), prevalences and mean scores are known for many patient groups.

It is beyond the scope of this review to give detailed figures for all the different patient groups, especially as prevalences have been reported for different cutoffs and therefore vary considerably. Nevertheless, there are some consistent findings that deserve mention.

Cross-sectional group differences. Women score higher on HADS anxiety than men [e.g., 3, 83, 86, 87, 96, 100, 128, 129, 132, 190], while gender differences are not significant in most studies for HADS depression. There is a nonlinear association of HADS scores with age [3]: HADS anxiety is highest in patients aged 30–59 years and lowest in those \( \geq 70 \) years. Depression scores are highest in 50–59-year-old patients and lowest in the youngest group (<30 years). Patients with poor psychosocial resources and low levels of information about their physical disease have higher anxiety [116, 132, 198] and especially depression scores [38, 93, 104, 132].

High HADS anxiety has been observed in patients with functional gastrointestinal [139, 141, 143, 146, 193], oral [192, 194], and cardiopulmonary [4, 28, 131, 132, 138, 153] symptoms. In patients scheduled for coronary angiography, high HADS anxiety discriminates patients with normal coronary arteries from those with coronary artery disease [4, 67]. Furthermore, HADS anxiety was high in patients with headache [148], genital herpes [174, 178], termination of pregnancy, or early miscarriage [43, 158, 160]. Although depression scores were also high in some of these groups, this pattern was clearly less convincing.

In contrast, HADS depression was more consistently associated with “passive” patient behavior such as noncompliance [154], low exercise tolerance [28, 132], or requests for general anesthesia in Cesarean section [162]. It was also associated with the activity of malignant [101, 196] or chronic [140] disease.

Similarly, patients with poorer performance due to cancer [104] or rheumatoid arthritis [29] and those with long-lasting angina pectoris [135, 138] are significantly more depressed than those with fewer symptoms. Severity of disease itself as measured by tumor site or size, presence of metastases [87, 101, 104], estrogen receptor status [105], electrocardiographic evidence of myocardial ischemia, number of coronary stenoses, left ventricular ejection fraction [28, 67, 132], medical prognosis [119], or proximity to death [104] is not positively related to HADS anxiety or depression scores. There are even reports of surprisingly low HADS scores in the most severely ill patients [e.g., 119, 132].

On the other hand, high anxiety and depression scores have been observed in patients with fibromyalgia, chronic fatigue [3, 149, 184, 187] or chronic pain [3, 148, 182], as well as in HIV-infected patients who choose to undergo nonconventional treatment [175]. In coronary patients, anxiety and depression scores are high after delivery of frequent shocks by an implanted defibrillator [134], in symptomatic (versus silent) myocardial ischemia [67]. And while awaiting by-
Table V.—Types of studies using HADS in different medical settings (figures represent individual studies as numbered in references list)

<table>
<thead>
<tr>
<th></th>
<th>A. Cross-sectional study types*</th>
<th>B. Longitudinal study types</th>
</tr>
</thead>
<tbody>
<tr>
<td>General (internal) medicine</td>
<td>1, 16, 17, 39, 41, 45, 51, 56, 59, 62, 64, 65, 83</td>
<td>6, 8, 84</td>
</tr>
<tr>
<td>Oncology</td>
<td>5, 11, 12, 18, 19, 47, 55, 57, 60, 63, 66, 87</td>
<td>7, 88, 89, 90, 91, 92</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>50</td>
<td>9, 50, 139, 140, 141, 142, 143, 144, 145</td>
</tr>
<tr>
<td>Rheumatology†</td>
<td>53</td>
<td>147</td>
</tr>
<tr>
<td>Category</td>
<td>Studies</td>
<td>152</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>Pneumology‡</td>
<td>43, 65</td>
<td></td>
</tr>
<tr>
<td>Gynecology/obstetrics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genitourinary medicine/HIV</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Psychiatry/psychosomatics</td>
<td>13, 27, 31, 40, 42, 54, 59, 179, 180, 181</td>
<td>3, 182, 183, 184</td>
</tr>
<tr>
<td>Other patient groups§</td>
<td>11, 34, 52, 65</td>
<td>3, 144, 174, 193, 194</td>
</tr>
<tr>
<td>Partners and caregivers</td>
<td>190, 191, 192</td>
<td>97, 198, 199</td>
</tr>
</tbody>
</table>

* Types of studies (see Table IV).
† Including fibromyalgia and back pain samples.
‡ Including assessment of respiratory symptoms in population surveys.
§ General (noncancer) surgery, nephrology/hemodialysis, dermatology, and oral medicine/dentistry.
### Table VI.—Experiences with the HADS outside the United Kingdom

<table>
<thead>
<tr>
<th>Country</th>
<th>Specific validation</th>
<th>Medical settings/study groups</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arab countries</td>
<td>++</td>
<td>Primary care, psychiatry, control subjects</td>
<td>6, 13, 16, 41</td>
</tr>
<tr>
<td>Australia</td>
<td>(+)</td>
<td>Stroke patients and caregivers, patients on a coronary care unit</td>
<td>20, 61, 78, 202, 209</td>
</tr>
<tr>
<td>Brazil</td>
<td>+</td>
<td>General medical wards</td>
<td>17</td>
</tr>
<tr>
<td>Canada</td>
<td>−</td>
<td>Rheumatoid arthritis, asthma,† closed head injury, abdominal bloating, Crohn’s disease, control subjects</td>
<td>29, 143, 156, 172</td>
</tr>
<tr>
<td>China/Hong Kong</td>
<td>++</td>
<td>General practice, diabetics, control subjects, medical students</td>
<td>15, 83, 145</td>
</tr>
<tr>
<td>Denmark</td>
<td>−</td>
<td>Asthma†</td>
<td>156</td>
</tr>
<tr>
<td>Finland</td>
<td>−</td>
<td>Asthma†</td>
<td>156</td>
</tr>
<tr>
<td>France/Belgium</td>
<td>++</td>
<td>Cardiologic and psychiatric outpatients, asthma,† medical in-patients, different groups of cancer patients</td>
<td>12, 39, 62, 63, 66, 68, 87, 120, 129, 130, 156, 186</td>
</tr>
<tr>
<td>Germany/Switzerland</td>
<td>++</td>
<td>Cardiology (exercise testing, coronary artery disease, implanted defibrillators), cancer, psychiatry, HIV, back pain, fibromyalgia, control subjects</td>
<td>3, 4, 10, 24, 67, 82, 112, 132, 133, 134, 135, 137, 175, 210</td>
</tr>
<tr>
<td>India</td>
<td>−</td>
<td>Psychiatry</td>
<td>26</td>
</tr>
<tr>
<td>Ireland</td>
<td>−</td>
<td>Asthma,† oral lichen planus</td>
<td>156, 192</td>
</tr>
<tr>
<td>Italy</td>
<td>−</td>
<td>Climacteric women, inflammatory bowel disease, asthma,†</td>
<td>142, 159, 156</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>(+)</td>
<td>Stroke, myocardial infarction, asthma,† control subjects</td>
<td>23, 156</td>
</tr>
<tr>
<td>Nigeria</td>
<td>+</td>
<td>Medical and surgical wards, gynecology and antenatal clinic, community sample</td>
<td>65</td>
</tr>
<tr>
<td>Norway</td>
<td>−</td>
<td>Lumbar disc surgery</td>
<td>173</td>
</tr>
<tr>
<td>Pakistan*</td>
<td>+</td>
<td>Tertiary care center, general practice, medical out-patient clinic, psychiatry, students</td>
<td>14, 38, 86, 179, 206, 211</td>
</tr>
<tr>
<td>Portugal</td>
<td>−</td>
<td>Asthma†</td>
<td>156</td>
</tr>
<tr>
<td>South Africa</td>
<td>(+)</td>
<td>Lung cancer</td>
<td>111, 156</td>
</tr>
<tr>
<td>Spain</td>
<td>(+)</td>
<td>Elderly persons</td>
<td>32</td>
</tr>
<tr>
<td>Sweden</td>
<td>+</td>
<td>Cancer (breast, lung, melanoma), asthma,† community sample, affections of the spinal cord, mothers of handicapped children, nursing staff</td>
<td>19, 25, 92, 94, 95, 96, 100, 105, 108, 109, 116, 153, 156, 167, 169, 199, 200, 201</td>
</tr>
<tr>
<td>Thailand</td>
<td>−</td>
<td>Asthma†</td>
<td>156</td>
</tr>
<tr>
<td>USA</td>
<td>−</td>
<td>Cancer in-and-out-patients, medical in-patients, asthma†</td>
<td>46, 101, 156</td>
</tr>
</tbody>
</table>

Specific validation data: −, not available; (+), minimal validation data; +, validation data available, but incomplete; ++, sufficient validation data.

* Urdu version of HADS, including Pakistani migrants in the UK.
† Subgroups of a multinational study.

pass surgery [136]. High preoperative HADS anxiety scores have also been reported for other patient groups [90, 197].

As could be expected, extremely high anxiety and depression scores have
been observed in psychiatric patients [e.g., 13, 31, 64, 69, 188], whereas lowest scores are reported for some groups of healthy subjects [3, 139, 141, 148, 158, 183]. Nevertheless, high HADS anxiety (and—to a lesser extent—depression) scores have also been observed in nonpatients, such as senior physicians [36], mothers of sick or psychotic children [198, 201], and spouses of patients suffering a first myocardial infarction [76].

**Cross-sectional correlations.** Cross-sectional correlation studies (Table V, Category A4) confirm the results of the group comparisons: In cancer patients, HADS depression scores are significantly correlated with performance status, pain, asthenia, and dyspnea [88, 96, 112]. Correlations of quality-of-life indicators with HADS depression or depression + anxiety sum scores reach r-values of up to 0.86 [25, 111]. HADS anxiety is inconsistently correlated with performance [88] and cancer pain [113]. In cardiac patients, HADS anxiety rather than HADS depression correlates with chest pain, tachycardia, dizziness, and total number of symptoms, whereas HADS depression correlates with dyspnea, (low) exercise tolerance, and most dimensions of a quality-of-life scale [24, 132, 134]. High correlations of HADS depression with quality of life have also been reported for patients with HIV infections, renal insufficiency, and spinal arteriovenous malformations [169, 177, 195]. Other significant correlations are found between HADS depression and pain (back pain [53], fibromyalgia [149], spinal arteriovenous malformations [169]), morning stiffness (rheumatoid arthritis [29]), dental plaque [22], sexual problems (patients with head injury and their partners [172]), harm avoidance (anxiety disorders [186]), and low self esteem (eating disorders [185]). HADS anxiety correlates with plasma cortisol, prolactin (gastrointestinal diseases [141]), somatic symptoms (lumbar disc prolapse [141]), vaginism (partners of patients with head injuries [172]), and visibility of burn injuries [191]. Both HADS subscales are correlated with the intensity of abdominal bloating [139]. In women awaiting surgery for breast cancer, they correlate with complaints about the quality of medical and psychological care [114].

**Longitudinal studies**

It is not surprising that there have been few longitudinal assessments of anxiety and depression. An exception is for cancer patients. One reason for this is that the HADS has been recommended as “a useful tool for measuring the psychological dimensions of quality-of-life in cancer patients” by the British Medical Research Council’s Cancer Therapy Committee Working Party on Quality of Life [208].

**Longitudinal changes in HADS scores. Influences of natural progress and somatic treatments in cancer patients.** While in the aforementioned cross-sectional studies the objective severity of physical disease was unrelated to HADS anxiety and depression, one quasilongitudinal study of patients with different stages of malignant melanoma showed a significant effect of tumor stage on HADS depression [19]. In a second, prospective study, the same investigators did not, however, find an increase of HADS scores over time nor a difference between groups with high- versus low-risk melanomas [100]. Several studies have assessed the effects of cancer treatment. Some of these
only mention that certain somatic therapies do not affect HADS scores and give little information about specific psychological processes [e.g., 119]. Others observed decreasing HADS scores during routine treatment [88, 103, 117, 118] as well as in the course of specific chemotherapies [96, 123], after surgery [72], bone marrow transplantation [37], or complementary cancer therapy [124]. However, it is not clear if these improvements simply reflect an alleviation of anticipatory anxiety, if they are due to adaptive coping processes, or to physical tumor response (pro [96]; contra [88]).

Most probably, tumor stage and response only have an indirect effect on subjective well-being, if patients become aware of them by changes in physical symptoms or by information obtained from their physicians. It should, however, be underscored that cancer patients are in general not abnormally anxious and only moderately depressed unless they have to face critical events or psychosocial problems [72, 102].

Spontaneous changes in noncancer patients. The ability of patients to cope with bad medical news has been demonstrated by some studies in patients with miscarriage [21], induced abortion [160], genital herpes [174], or a new diagnosis of HIV infection [58]: these patients showed a rapid decrease of initially high HADS scores. However, another group of women with early miscarriage [158] did not experience substantial alleviation of anxiety, and, in the HIV study, patients that screened negative had persistently high HADS scores which might be due to their continuing risk of acquiring the virus. Increasing HADS anxiety scores were observed at a 2-year follow-up in patients with nonspecific abdominal pain [9]. HADS depression scores increase significantly during pregnancy, with smoking as a significant univariate predictor [161].

Effects of somatic treatments in noncancer patients. There have also been a number of longitudinal studies to evaluate medical or surgical interventions in noncancer patients and healthy controls. Some drugs (e.g., modern cholesterol-lowering agents or indomethacin) were tested in healthy persons to examine possible influences on mood and cognitive performance, but HADS scores were not changed by these substances [205, 207]. In patients with rheumatoid arthritis, out-patient instead of in-patient treatment did not adversely affect HADS anxiety or depression [150], neither did a treatment with methotrexate and folic acid [151]. In a randomized study of nedocromil sodium in asthma patients, HADS scores decreased significantly in the intervention and control groups with only marginal advantages for the verum group [156]. In gynecological patients, no differences were observed between HADS scores after medical versus surgical abortion [160]. Women with premenstrual tension syndromes had good relief of symptoms with moderate but significant decreases of HADS anxiety and depression during the first two treatment cycles with goserelin [163]. Treatment with aciclovir for recurrent genital herpes went along with decreasing HADS anxiety [178]. In general surgical patients, HADS anxiety (and somewhat less depression) scores decreased significantly after total knee replacement [197]. Once more, this might also be due to the disappearance of preoperative anxiety. In a group of cardiological patients, we did not find a beneficial effect of bypass surgery or coronary angioplasty on psychosocial outcome when baseline HADS scores were obtained before the decision for the intervention
was made. Instead, this sample with severe coronary disease experienced an increase of HADS scores over time which was independent of cardiological baseline findings or treatment and could mainly be explained by chronic disability and adverse life events [24].

Effects of psychosocial intervention. Given the high rate of spontaneous remission, the aforementioned studies (see Treatment validation subsection) of psychosocial interventions have to be re-examined. Indeed, a part of the effect ascribed to psychotherapy might also have occurred with routine treatment only. Thus, in one study there was no difference in the course of HADS anxiety and depression between two groups of cancer patients with psychotherapy or a placebo chat, and both groups had only small, transient improvements in anxiety, but not depression, as compared with a conventionally treated control group [72]. In another controlled intervention study in cancer patients, the greatest decrease of significant anxiety was observed between a baseline screening and the start of the intervention. Nevertheless, at follow-up there was a persisting difference in favor of the psychotherapy group whose anxiety was once more significantly reduced by the intervention [71]. Significant advantages of psychosocial treatment over routine care have also been documented in controlled intervention studies with cardiological and psychiatric patients (see above).

Prediction of psychosocial and physical outcome. Prediction of outcome by means of baseline HADS scores has been attempted in several studies (Table V, Category B3): Some investigators simply predicted follow-up HADS scores in cancer patients by means of baseline HADS scores [70, 72, 121]. Given the good retest reliability of the HADS, this prediction is not surprising. However, it is worth mentioning that, in one of these studies [70], none of several disease-related variables significantly predicted HADS scores in multivariate regression. HADS global scores also predicted anticipatory nausea in patients undergoing adjuvant chemotherapy for breast cancer [120].

In renal transplant patients, baseline HADS anxiety and depression scores predicted follow-up psychiatric morbidity as assessed by standardized interview. In addition, high pretransplant HADS anxiety scores predicted posttransplant referral to a psychiatrist [196]. In a group of patients with coronary artery disease baseline HADS depression scores were the best predictors for most dimensions of quality of life after 2.6 years [24].

Physical outcome was successfully predicted in three studies: psychiatric “caseness” according to the HADS predicted bad remission in patients with Crohn’s disease [50]. HADS anxiety and—less clearly—depression predicted outcome in patients undergoing surgical treatment for chronic constipation or fecal incontinence [193]. In lumbar disc surgery, clinical outcome was significantly predicted by HADS anxiety in a multivariate model [173].

Prediction of mortality was attempted in patients with advanced lung cancer [127] or bone marrow transplantation [37]. However, none of the psychological predictors including HADS scores was significant in these studies. Currently, we are undertaking a longitudinal assessment of 455 consecutive medical inpatients. Preliminary, still unpublished results indicate that HADS scores at admission do not predict mortality in cancer patients, whereas in noncancer pa-
tients high HADS depression is a significant predictor of 1-year mortality, even after statistically controlling for severity of symptoms, disability, and comorbidity.

Experiences with the HADS outside the UK

The publisher of the HADS provides at least 33 translations of the scale. However, many of them have not been validated in their national and cultural environments. On the other hand, some investigators have created and validated their own translations and used them in different clinical settings. Empirical data are available from about 25 countries outside the United Kingdom. The scale can be considered sufficiently validated for use in Arab countries, China, France and Belgium, and Germany and Switzerland (Table VI). For several other countries, only partial validity information is available.

Some national versions have been used successfully in clinical studies but their formal characteristics are insufficiently documented. Thus, they may give useful information in group comparisons or longitudinal studies, but it is unknown whether they are valid and reliable instruments for individual screening. It is also not known if these versions are equivalent to the English version, or if their scores are influenced by cultural factors. In a large multinational treatment study with asthma patients [156], nationality significantly influenced the association between HADS scores and a single item asking for general health. The investigators in that study believe this to be due to cultural or linguistic influences on the different national HADS scores. However, it is also possible that the unvalidated control item might be culturally influenced or that the relevance of anxiety and depression for perceiving general health varies across countries. Sufficient validity data for the different national HADS versions could help to answer this question. For example, there is some evidence that the Urdu version of the HADS is rather similar to the English one when both scales are completed by the same bilingual persons [14]. However, a comparative study using both versions with British and Pakistani attenders of two English general practices showed significantly more depression and a significantly higher correlation of both HADS subscales with somatic symptoms in the migrant group [38]. This indicates that, despite probable identity of psychometric properties, HADS scores may be different in countries with different cultural patterns of perceiving and expressing emotions, which is an important issue when transferring the scale to new cultural settings. It would be especially worthwhile to evaluate the very well-validated English HADS version in the United States, because there is still a lack of specifically designed questionnaires for assessing anxiety and depression in American medical patients. Experiences from Canada and the United States itself (see Table VI) are very promising and justify further investigation.

CONCLUSIONS

The HADS is a reliable and valid instrument for assessing anxiety and depression in medical patients. Its construction facilitates its use with these patients and it could be expected to have considerable advantages over other established instruments. However, this theoretical advantage remains to be proven empiri-
cally. While available data support some of the theoretical assumptions (e.g., the two-dimensional scale structure, its sensitivity for mild symptoms and changes over time, or its good acceptance), some original concepts should be abandoned: In contrast to earlier assumptions, the scale does not identify a specific form of drug-responsive or “biogenic” depression and does also not allow one to make a diagnosis of major depression. The concept of anhedonia which underlies the depression subscale is a nonspecific indicator of depressed mood [180], but this is, in my opinion, a strength of the scale, because this is what makes it sensitive to mild disturbances without relying on somatic symptoms. Because many investigators use the HADS in a rather unreflective way one can be happy that high HADS scores do not simply justify antidepressant drug treatment. Although HADS scores are responsive to drug treatment, more convincing data suggest their responsiveness to psychosocial intervention. Thus, in clinical practice, high HADS scores with respect to a specified cutoff or other reference value, should lead to a thorough psychiatric interview in which a definite diagnosis can be made and an adequate treatment initiated.

For scientific purposes, the scale is able to differentiate groups with different prevalences or intensities of anxiety and depression. It allows longitudinal assessments with repeated testing at intervals of about 1 week or more and is sensitive to changes in patients’ emotional state. It is well documented to predict mood over intervals of 1 year and longer. It also predicts compliance, quality of life (HADS depression), and physical symptoms (HADS anxiety). Whether the HADS also predicts objective somatic endpoints still remains uncertain. However, preliminary data suggest that activity of Crohn’s disease and mortality of medical noncancer patients might be predicted by baseline HADS scores.

Given the high subscale correlation, it has been a matter of discussion whether the HADS should still be used as a two-dimensional instrument or as a single measure of emotional distress. Looking at the data, such a global measure may be adequate if detailed analysis would not result in additional consequences, for instance, in studies of chemotherapy in cancer patients. On the other hand, both subscales have somewhat different clinical characteristics: The HADS anxiety subscale is mainly responsive to “active” patients with “passive” disease, such as those complaining of multiple or functional symptoms with stable or absent organic disease. In contrast, the HADS depression subscale is cross-sectionally and longitudinally associated with “passive” patients and “active” disease. This means little patient cooperation, such as compliance or exercise endurance, and active malignant or chronic disease and disability. Although there is some overlap, both subscales should therefore be used separately for most purposes.

The opportunity afforded by the multiple translations of the HADS should be taken to use the scale in multinational studies. This could also help to establish context-specific reference values. Finally, ongoing international experience might help to improve the instrument itself. An international consensus group or online forum might be beneficial for gathering national data and developing revised international versions of the HADS which might then become an international research standard.
REFERENCES


Review

36 Review


83. Lam CL, Pan PC, Chan AW, Chan SY, Munro C. Can the Hospital Anxiety and Depression (HAD) scale be used on Chinese elderly in general practice? Fam Pract 1995;12:149–154.

84. Walker LG, Cordiner CM, Gilbert FJ, Needham G, Deans HE, Affleck IR, Hood DB, Mathieson


139. Heaton KW, Ghosh S, Braddon FEM. How bad are the symptoms and bowel dysfunction of patients with the irritable bowel syndrome? A prospective, controlled study with emphasis on stool form. Gut 1991;32:73–79.


