

NOVA University of Newcastle Research Online

nova.newcastle.edu.au

Carey, Mariko L.; Noble, Natasha; Sanson-Fisher, Robert; Mackenzie, Lisa, 'Identifying psychological morbidity among people with cancer using the Hospital Anxiety and Depression Scale: time to revisit first principles?', Psycho-Oncology Vol. 21, Issue 3, p. 229-238 (2012)

Available from: http://dx.doi.org/10.1002/pon.2057

Accessed from: http://hdl.handle.net/1959.13/1042071

Identifying psychological morbidity among people with cancer using the Hospital Anxiety and Depression Scale: Time to revisit first principles?

Mariko Carey, Natasha Noble, Robert Sanson-Fisher, Lisa Mackenzie

Priority Research Centre for Health Behaviour

Faculty of Health

The University of Newcastle

New South Wales,

Australia

Corresponding author:

Dr. Mariko Carey

Address: Room 266 David Maddison Building,

The University of Newcastle

University Drive Callaghan NSW Australia 2308

Telephone: (61 2) 4913 8320

Fax: (61 2) 4913 8779

Email: Mariko.Carey@newcastle.edu.au

Acknowledgements:

This research was supported by a National Health and Medical Research Institute Grant (grant no. 1010536). Dr. Carey is supported by Hunter Medical Research Institute Post-Doctoral Fellowship.

Word count: 3,611

Key words: cancer, oncology, psychological distress, depression, anxiety, screening

Abstract

Background: The aim of this review was to describe the findings and methodological quality of studies which sought to validate the Hospital Anxiety and Depression Scale (HADS) against the Structured Clinical Interview for DSM (SCID) in cancer populations. We also sought to compare the cut points recommended by these validation studies to the way in which the HADS is currently used to determine prevalence of psychological morbidity in cancer populations.

Methods: An electronic database search was conducted of Medline from 1983 to October 2010 for validation studies of the HADS in cancer populations. Reference lists of HADS reviews were hand searched. To examine which cut points are commonly used in cancer specific literature to identify the prevalence of psychological disorders, studies published in 2009 were identified via an electronic database search of Medline.

Results: 10 studies which validated the HADS against the SCID in cancer patient populations were found and examined in detail. None met all methodological criteria associated with the selection of a screening instrument. Recommendations for optimal HADS thresholds varied substantially across these studies. The most commonly used threshold for determining caseness in the 2009 literature on prevalence of psychological distress among cancer patients was a subscale score of \geq 8. This threshold was poorly supported by the results of the 10 cancer HADS validation studies examined.

Conclusions: Caution is warranted in interpreting the results of prevalence studies using the HADS. There is a need to develop evidence about the optimal thresholds for defining caseness using the HADS.

Many studies have reported high rates of psychological morbidity among cancer patients [1-3]. This has led to growing efforts to improve psychosocial care for people with cancer. In particular, there has been an increasing emphasis on finding effective and cost efficient ways to identify those at risk of poor psychological outcomes, and providing them with psychological therapies suited to their needs [1, 3].

The Hospital Anxiety and Depression Scale (HADS) is frequently used in establishing prevalence of psychological distress in cancer patient populations [4-6]. The HADS consists of 14 items which can be used to assess anxiety (7 items) and depression (7 items), to provide a general indication of distress. Each HADS item is scored from 0 to 3, giving a maximum score of 21 for each subscale. The HADS developers initially suggested that subscale scores of less than 7 indicated non-cases of anxiety or depression, scores of 8-10 'doubtful-cases' <u>'possible' or 'borderline' cases</u>, and scores of 11 or more <u>'probable' or</u> 'definite' cases' [7]. In this original study of just 100 medical patients, the HADS was validated against a 20 minute 'psychiatric assessment' rather than a standardised interview tool [7]. Sensitivity and specificity were not reported, [7], suggesting the need for rigorous validation studies before widespread use of the instrument. Later the authors provided recommended interpretations for HADS scores with subscale scores of 0-7 indicating normal, 8-10 indicating mild, 11-14 indicating moderate, and 15-21 indicating severe, anxiety or depression [8]. However, these recommendations were not accompanied by supporting data detailing the sensitivity and specificity of these thresholds [8]. Factor analysis has indicated, however, that the instrument assesses distinct constructs of anxiety and depression [9].

Compared with other patient-reported outcome measures for assessing depression, the HADS has less emphasis on the somatic features of depression such as fatigue, sleep disturbance and appetite [10, 11]. It has been argued that the omission of these features from the items of the instrument make it more suitable than other measures for assessing psychological well-being in medically ill populations, where such depressive symptoms may be confounded with symptoms of the disease [4]. As a consequence, the HADS has become one of the most widely used tools for assessing psychological morbidity in cancer patients [4, 10, 12, 13]. Several review papers have suggested that it is among the best tools for assessing psychological outcomes for this population [4, 11, 13].

There are two main ways in which the HADS can be used to examine psychological morbidity in people with cancer. The first involves calculating a mean score and standard deviation [4]. Mean scores can be compared for different groups of patients or over time. While statistical definitions of what constitutes a significant change can be applied to mean scores [14, 15], there is debate about how well such changes reflect clinically significant changes [15].

The second way of using the HADS to examine psychological morbidity involves the use of a threshold score or range of scores representing a cut point. Because it is often necessary to make a judgement about which people are in need of professional help to address the distress detected using the HADS, thresholds are commonly used to determine caseness instead of a mean score [4, 16, 17]. Such thresholds are ideally determined on the basis of optimising sensitivity and specificity compared with a gold standard diagnostic test, which

indicates that the threshold score is clinically meaningful [4, 10, 17]. Therefore the use of thresholds may have significant advantages over the use of mean scores.

When thresholds are used to identify those with clinically significant distress, however, the sensitivity and specificity of the chosen threshold needs to be taken into account. Sensitivity relates to how well a test detects all those with the condition of interest compared to a gold standard test, while specificity refers to how likely the test is to identify only those with the condition of interest [18]. Thus a highly sensitive test will have a low false negative rate; and a highly specific test will have a low false positive rate [18]. The level at which these thresholds for defining caseness are set has important implications for establishing the prevalence of psychological morbidity among cancer patients.

While there is no biochemical or pathological test which acts as a gold standard for assessing the presence of a psychiatric disorder [19], the Structured Clinical Interview for DSM (SCID) is currently used to diagnose Axis I psychological disorders as defined by the Diagnostic and Statistical Manual of Mental Disorders [20] (DSM; versions III, IV and IV-TR). Other structured diagnostic interviews, such as the Schedule for Affective Disorders and Schizophrenia (SADS), do not include diagnoses for all the major Axis I disorders included in the DSM [20]. The SCID has been periodically updated to reflect revisions in the diagnostic criteria of successive versions of the DSM and includes diagnostic criteria taken directly from the DSM. Internationally, the DSM is the most commonly used and valued diagnostic tool for psychiatric research purposes [21]. Several bodies also recommend that 'distress' among cancer patients be operationalised as adjustment disorder or another psychiatric disorder from the DSM [22]. The SCID may thus be considered the gold standard comparator for an instrument such as the HADS.

Therefore, it is timely to critically examine which thresholds are being used to establish the prevalence of psychological morbidity among cancer patients, as well as to consider the methodological quality of the validation studies supporting the use of particular thresholds. This review aimed to: 1) Describe the findings and methodological quality of the studies which sought to validate the HADS against the SCID in cancer populations; and 2) Compare the recommendations for thresholds/cut points arising from validation studies to the way in which the HADS is currently used in the cancer literature.

Methods

Literature search to identify validation studies

The reference lists of several recent systematic reviews of the HADS specifically [10, 11], or of distress screening instruments in general [23], were hand searched to identify studies in which the HADS was validated against the SCID. In addition to this, an electronic database search of Medline was conducted to identify relevant studies published between 1983 (the year in which the HADS was first published [7]) and 30th October 2010 which may have been missed by previous reviews. The following search terms were used: *Hospital Anxiety and Depression Scale* or *HADS, cancer* or *neoplasm,* and *validation studies* or *psychometric properties* or *interview* or *clinical interview*. Only studies which reported an optimal threshold on the HADS in comparison to the SCID were included.

Methodological quality of studies

Validation studies were assessed against the following five criteria which are important considerations when selecting a screening test [24-26]. Data were extracted <u>independently</u> by two authors and then cross checked to confirm accuracy. by one author and checked by another.

Criterion 1: Appropriateness of sample. The sample in which the validation study takes place is important for determining the appropriateness of the test for use in other populations. This is because the predictive value of the test will change with the prevalence of the condition of interest [24, 27]. The sample should also include an appropriate spectrum of the disease of interest [26]. Studies were coded as to whether they provided an adequate description of the sample. Age, gender, sample size, break down of cancer type/s and the setting from which the sample were recruited were considered minimum requirements.

Criterion 2: Precision estimates for the test result. Sensitivity and specificity values derived from empirical studies are estimates of the true value of these properties of the test [28]. Therefore it is important that confidence intervals are reported so that the range of values in which the true value lies can be observed by the reader [28].

Criterion 3: Independent and blind assessment. This refers to whether the individual/s who conducted the clinical interview were "blind" or unaware of the person's scores on the screening or diagnostic test when they conducted the interview. This reduces the risk of interviewer bias affecting the results [26].

Criterion 4: Reliability of "gold standard" clinical interview. For a clinical interview to be used as a gold standard assessment of psychological morbidity, it must be reliable. That is, it must have test-retest reliability and inter-rater reliability [29]. Test-retest reliability is a property of the assessment tool itself and need only be demonstrated once. However, interrater reliability may vary according to the training and skills of the interviewers [30], and therefore should be reported in validation studies. Generally an inter-rater agreement of 90% or kappa \geq 0.80 is considered adequate [31].

Criterion 5: Utility of the test with respect to improving patient outcomes. The utility of the test lies in whether use of the test ultimately results in better outcomes for those tested. This is reliant on the availability of effective treatments for those identified as cases [26].

Literature search to identify use of HADS thresholds in the literature

A search of Medline using the terms *Hospital Anxiety and Depression Scale* or *HADS* and *cancer* or *neoplasm*, was conducted for 2009 (January-December inclusive). This search was intended to provide a snap shot of how the HADS has recently been used in the cancer literature rather than a comprehensive assessment. Any studies which assessed the prevalence of psychological distress among cancer patients using the HADS were included. The threshold score on the HADS which was used by the authors to determine caseness, as well diagnostic outcomes considered (e.g. depression, any disorder etc.) was extracted for each study.

Results

HADS validation studies

Comment [m1]: Tash can you add details of the PsychInfo search here?

A total of 11 validation studies were identified that compared the performance of the HADS against the SCID. However, two of these studies used the same sample of cancer patients for validation of the HADS and recommended the same optimal HADS cut point [32, 33]. The results of these two studies are therefore only reported once in the analysis below.

The majority of the validation studies assessed patients for SCID diagnoses of adjustment disorder and/or major depressive disorder. A number of studies, however, gave patients a diagnosis of 'any psychiatric disorder', which in some samples included anxiety disorders, alcohol dependence and psychosis. Five studies examined sensitivity and specificity separately for more than one diagnostic category. Six studies examined major depression; five adjustment disorder with or without depression; and four examined any psychiatric disorder.

Methodological quality of studies validating HADS against the SCID

Due to the inclusion criteria of the review, all studies met criterion one, as all reported on the validation of HADS in a cancer population. All studies also reported sensitivity and specificity, however, only two of the 10 studies met criterion two by reporting precision estimates (confidence intervals) for sensitivity and specificity [34, 35]. Four studies reported that interviewers were blind to participants' HADS scores (criterion three; [33, 35-37]), and of these, only one also met the criterion related to precision estimates [35]. Only two studies reported on the inter-rater reliability of the clinical interview (criterion four). Neither of these two studies met the reliability criteria for all diagnoses assessed within the study (an inter-rater agreement of 90% or kappa \geq 0.80). Akechi *et al.* (2006) [35] met inter-rater reliability criteria for major depression but not adjustment disorder; Keller *et al.* (2004) [38] met inter-rater reliability criteria for any psychiatric diagnosis, but not for specific diagnoses. No studies reported on the clinical utility of HADS with respect to patient outcomes (criterion five). Figure 1 displays the quality criteria in a hierarchy, showing the number of studies meeting each successive criterion as well as all the preceding criteria. Details of each validation study, including the recommended cut points and sensitivity and specificity are presented in Table 1.

<INSERT FIGURE 1 HERE>

<INSERT TABLE 1 HERE>

Studies using the HADS to assess prevalence of psychological distress

The literature search for articles using a HADS threshold score to assess prevalence of psychological distress in cancer populations identified 79 studies published in 2009. Of these, 55 articles were excluded: 6 articles were not published in English, 10 were duplicates, 5 reported on the use of HADS in non cancer samples, 3 were review articles, 3 were validation studies, 4 were not relevant, 21 were prevalence studies reporting mean scores or correlations rather than using thresholds, and in 3 studies, the HADS threshold scores used were not stated. This left 24 relevant prevalence studies which reported on the use of HADS threshold scores.

These 24 prevalence studies used the HADS to determine variously cases of 'psychological or emotional distress', 'psychological morbidity', 'elevated levels of anxiety or depression' or 'cases of depression' only. The HADS cut points used to identify distress or elevated anxiety/depression are presented in Table 2 below. Three of the prevalence studies used the HADS to determine cases of depression only, with these cut points shown in Table 3. These cut points were compared to the threshold scores determined by the validation studies above in order to maximise the sensitivity and specificity of the HADS as compared to the gold standard measure the SCID. Table 2 shows cut points used to assess prevalence compared to the optimal HADS threshold calculated for a SCID diagnosis of any psychiatric disorder or adjustment disorder; while Table 3 shows prevalence cut points in comparison with the optimal HADS threshold calculated for a SCID diagnosis of Major Depressive Disorder.

<INSERT TABLE 2 HERE>

<INSERT TABLE 3 HERE>

As shown in Tables $\underline{23}$ and $\underline{34}$, there was a poor correspondence between the thresholds used to assess prevalence of distress and those thresholds established in the validation studies in order to maximise sensitivity and specificity of the HADS relative to clinical interview. Thresholds to determine caseness vary substantially in the validation literature, for example from a subscale score as low as 4 [39] to as high as 11 [40]. In contrast, the most commonly used HADS threshold to determine prevalence was a subscale score of 8 or above to identify a 'case' of anxiety, depression or distress, used in 13 of the 24 prevalence studies. This score (\geq 8) was only recommended in one of the ten relevant validation studies [40]. Several prevalence studies used thresholds which were not supported by any of the cancer validation studies identified in this review.

Discussion

Of the 10 studies identified in which the HADS was validated against the SCID in cancer populations, none met all the methodological criteria required of screening and diagnostic tests. A number of methodological problems are evident in the validation studies identified. Confidence intervals associated with sensitivity and specificity calculations were poorly reported. Without such information, it is difficult to determine where these true values lie [28]. For the two studies which reported precision estimates, a lack of precision in the reported sensitivity and specificity of the thresholds was notable. This most likely reflects relatively small samples used in the validation studies and low prevalence of the disorders being examined. Criteria related to blinding and inter-rater reliability for the clinical interview were also poorly met across studies. The reliability of any intervieweradministered instrument is a function of many factors, including interviewer training and the characteristics of the subject sample [30]. Therefore, given the limited reporting of blinding and inter-rater reliability, it is possible that these factors adversely affected the accuracy of the sensitivity and specificity estimates reported. It is notable that no studies reported on whether the use of the screening test led to better patient outcomes. This is an important issue to consider when weighing up the cost effectiveness of screening of any type [41, 42], and also raises ethical issues regarding whether the instrument should be used as a screening tool in the absence of any evidence of benefit. A recent review of randomised trials of the effect of screening for psychological distress on psychological outcomes for cancer patients, for example, reported a limited effect of screening [22]. Only three of seven randomised studies showed an effect of screening on psychological wellbeing [22].

13

What might account for variability in HADS thresholds reported in cancer validations studies?

The variability in the thresholds recommended in the HADS validations studies was notable. Apart from the methodological differences there are several other factors related to sample composition which may account for the variation in the HADS cut-points recommended in the cancer validation studies.

Variation in disease type and prognosis: Although the HADS was developed and validated in a medical population [7], it was not specifically developed for use with patients with cancer. Cancer is <u>a</u> heterogeneous disease, with different cancer types associated with varied prognostic expectations. For example, overall 5 year survival rates are more than 85% for breast and prostate cancer, but less than 16% for pancreatic and lung cancers [43, 44]. Different disease types and stages are also associated with different symptoms, treatment courses, and severity and endurance of treatment side effects [45-47]. Hence at particular stages of the disease trajectory and under particular circumstances a degree of worry or negative affect may be considered part of a normal, if unpleasant, adjustment to the clinical course of the disease [48, 49]. These differences need to be taken into account when interpreting the validity of recommended HADS thresholds across patients with different disease and treatment regimens [50].

Possible cultural effects on the thresholds recommended by validation studies: Of the 10 cancer validation studies identified for the HADS using the SCID as a gold standard, only one was conducted using the English language version of the HADS. Validation studies of

different language versions of the HADS have been associated with different factor structures [51] and optimal thresholds for identifying caseness [10]. It has been suggested that HADS thresholds may vary cross culturally as a result of variations in the symptomatic presentation of anxiety and depression [11, 52, 53]. For example, it has been suggested that culture may influence whether depression is expressed in emotional and psychological terms or whether it is manifested as physical symptoms [54]. As different items within the HADS focus either on physical or psychological symptoms, endorsement of different combinations of items in a given population will alter the specificity and sensitivity of a given threshold for defining caseness [55]. Therefore one might expect that the threshold for defining caseness may vary between cultures depending on the way in which cultural norms influence respondents' answers.

How well do thresholds used in prevalence studies correspond to the recommendations from validation studies?

While the present study presented only a 'snap shot' for the way in which the HADS thresholds are used to assess prevalence of psychological morbidity in the research literature, it nonetheless highlights important issues which warrant consideration. As shown in Tables 3 and 4, on both the anxiety and depression subscales, the most commonly used subscale threshold used in prevalence studies is a score of 8 or more. This threshold is poorly supported by validation studies with cancer patients, but does correspond to the recommendations <u>of</u> the original validation study by Zigmond & Snaith [7], conducted with a sample of just 100 medical patients. <u>As described previously, In this original study</u>, the HADS was <u>originally</u> validated against a 20 minute 'psychiatric assessment' rather than a standardised interview tool. Sensitivity and specificity were not reported [7]. These methodological limitations suggest the recommended thresholds from this original study may not be optimal for use with cancer populations.

The continued reliance on the Zigmond & Snaith thresholds [7, 8] may reflect short comings in the cancer validation literature. Validation studies conducted with HADS in cancer populations offer little consistency with respect to the thresholds; this is possibly due to variability within the cancer patient populations in terms of culture, disease stage, treatment status, and type of disease across the studies. The validation of the performance of the HADS against different SCID diagnoses may also have contributed to this variability. Psychological 'distress' is poorly operationalised and it is not entirely clear how distress, anxiety or depression correspond to the psychiatric diagnoses derived from the SCID. These factors have significant implications for the way that HADS is used and interpreted in psycho-oncology research with respect to estimating prevalence of psychological morbidity, defining treatment effects and the deployment of resources for psychosocial care.

Implications of use of HADS thresholds which may not be appropriate for the target cancer population

Potential error in prevalence estimations: The wide variation in the recommended thresholds across studies has significant implications for interpretation of the prevalence of psychological morbidity reported in the literature. Depending on the threshold used, this may result in a significant under or over estimation of psychological morbidity.

Implications for triage and access to psychosocial care: <u>While people with cancer may</u> require psychosocial services such as counselling for a variety of reasons [56] <u>symptoms of</u> anxiety and depression may be common triggers for referral. <u>This has resulted in increasing</u> emphasis on the use of HADS and similar measures as screening tools in clinical settings [13,

57].If thresholds are set too low when the HADS is used as a screening instrument, then it is likely that many patients who do not need specialist psychosocial care will be offered it. This may have several implications for the service. First, resources may be expended unnecessarily on patients who do not need specialist psychosocial care. Further, uptake of referral and other psychosocial services are likely to be lower than expected if patients are offered services which they do not feel that they need [1, 58]. This may lead to the appearance of low demand for such services, potentially resulting in funding cuts. Conversely, if thresholds are set too high, many patients who need psychosocial care may not be identified and offered appropriate care [38]. Untreated anxiety and depression are associated with detrimental outcomes such as poor quality of life, poor adherence to treatment, reduced recall of medical information, and poorer overall adjustment [59-61].

Estimates of treatment effectiveness: HADS is often used as an outcome measure in psychosocial intervention trials in cancer populations [4]. If the approach of examining change in the proportion of people above the threshold for caseness is used, a threshold which is set too low may fail to detect an intervention effect. Conversely if the threshold is set too high, this may not capture change in those people with clinically significant distress who fall below the threshold. This may result in underestimation of the treatment effect. With continuing debate over which psychosocial interventions are effective and with which

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Formatted: Font: Not Bold

populations [62, 63], potential inaccuracies with respect to outcome measures add further complexity to interpreting this literature.

Conclusions

Examination of the current literature suggests that results of prevalence studies using the HADS should be interpreted with caution. There is also an urgent need to develop consistent evidence about how the HADS should be used in different oncology settings and what thresholds may be appropriate for identifying the prevalence of clinically significant psychological morbidity. This should be done through rigorous validation studies in defined cancer populations. These studies should clearly articulate the sample, and methods, including the type of disorder that the HADS is being used to identify. Clear and accurate reporting of these issues will assist readers in making judgements about the appropriate use of this measure in a given population.

References

- 1. Carlson LE, Angen M, Cullum J *et al.* High levels of untreated distress and fatigue in cancer patients. *Brit J Cancer* 2004; **90** (12): 2297-2304.
- 2. Spiegel D. Cancer and depression. Brit J Psychiat 1996; 168 (30 (Suppl)): 109-116.
- 3. Zabora J, Brintzenhofeszoc K, Curbow B *et al*. The prevalence of psychological distress by cancer site. *Psycho-Oncology* 2001; **10**: 19-28.
- 4. Luckett T, Butow P, King M *et al.* A review and recommendations for optimal outcome measures of anxiety, depression and general distress in studies evaluating psychosocial interventions for English-speaking adults with heterogeneous cancer diagnoses. *Support Care Cancer* 2010; **18** (10): 1241-1262.
- 5. Mitchell AJ. How accurate is the HADS as a screen for emotional complications of cancer? A metaanalysis against robust psychiatric interviews. *Psycho-oncology* 2009; **18** (S2): S86.
- 6. Wasteson E, Breenne E, Higginson IJ *et al.* Depression assessment and classification in palliative cancer patients: a systematic literature review. *Palliative Med* 2009; **23** (8): 739-53.
- Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiat Scand 1983;
 67 (6): 361-370.
- 8. Zigmond AS, Snaith RP. eds. *The Hospital Anxiety and Depression Scale with the Irritability Depression-Anxiety Scale and the Leads Situational Anxiety Scale Manual*. 1994; GL Assessment Ltd: London.
- 9. Moorey S, Greer S, Watson M *et al.* The factor structure and factor stability of the hospital anxiety and depression scale in patients with cancer. *B J Psych* 1991; **158**: 255-259.
- 10. Bjelland I, Dahl AA, Haug TT *et al.* The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* 2002; **52** (2): 69-77.
- 11. Herrmann C. International experiences with the Hospital Anxiety and Depression Scale-A review of validation data and clinical results. *J Psychosom Res* 1997; **42** (1): 17-41.
- Mehnert A,Koch U. Psychosocial care of cancer patients international differences in definition, healthcare structures, and therapeutic approaches. *Support Care Cancer* 2005; 13 (8): 579-588.
- Ziegler L, Hill K, Neilly L *et al.* Identifying psychological distress at key stages of the cancer illness trajectory: A systematic review of validated self-report measures. *J Pain Symptom Manag* 2011; **41** (3): 619-36.
- 14. Farivar S, Liu H, Hays R. Half standard deviation estimate of the minimally important difference in HRQOL scores? *Expert Rev Pharmacoecon Outcomes Res* 2004; **4** (5): 515-523.
- 15. Cook CE. Clinimetrics Corner: The Minimal Clinically Important Change Score (MCID): A Necessary Pretense. *J Man Manip Ther* 2008; **16** (4): E82-E83.
- 16. Crawford JR, Henry JD, Crombie C *et al*. Brief report: normative data for the HADS from a large non-clinical sample. *Br J Clin Psychol* 2001; **40**: 429-434.
- 17. Singer S, Gõtze H, Mõbius C *et al*. Quality of care and emotional support from the inpatient cancer patient's perspective. *Langenbeck Arch Surg* 2009; **394** (4): 723-31.
- 18. Fletcher RH, Fletcher SW. *Clinical epidemiology: The essentials*. Lippincott Williams and Wilkins: Baltimore, 2005.
- 19. Faraone SV,Tsuang MT. Measuring diagnostic accuracy in the absence of a "Gold Standard". *Am J Psychiat* 1994; **151** (5): 650.
- 20. Spitzer RL, Williams JBW, Gibbon M *et al.* The structured clinical interview for DSM-III-R (SCID) I: History, rationale, and description. *Arch Gen Psychiatry* 1992; **49**: 624-629.
- 21. Mezzick JE. International surveys on the use of ICD-10 and related diagnostic systems. *Psychopathology* 2002; **35**: 72-75.
- 22. Bidstrup PE, Johansen C, Mitchell AJ. Screening for cancer-related distress: summary of evidence from tools to programs. *Acta Oncol* 2011; **50** (2): 194-204.

- 23. Vodermaier A, Linden W,Siu C. Screening for Emotional Distress in Cancer Patients: A Systematic Review of Assessment Instruments. *J Natl Cancer I* 2009; **101** (21): 1464-1488.
- 24. Sackett DL, Haynes RB, Tugwell P. *Clinical Epidemiology: A Basic Science for Clinical Medicine*.Little, Brown and Company: Boston/Toronto, 1985.
- 25. Haynes RB, Sackett DL, Guyatt G *et al. Clinical Epidemiology: How to do Clinical Practice Research*.Lippincott Williams & Wilkins: Philadelphia, 2006.
- 26. Sackett DL, Haynes RB, Guyatt G *et al. Clinical Epidemiology. A Basic Science for Clinical Medicine*. Second Edition ed.,Little, Brown and Co: Boston, 1991.
- 27. Gerstman B. *Epidemiology Kept Simple. An Introduction to Traditional and Modern Epidemiology.*Wiley-Liss Inc: New Jersey, 2003.
- 28. Harper R,Reeves B. Reporting of precision of estimates for eiagnostic accuracy: a review. *Brit Med J* 1999; **318** (7194): 1322-1323.
- 29. Segal DL, Hersen M, Van Hasselt VB. Reliability of the structured clinical interview for DSM-III-R: An evaluative review. *Compr Psychiat* 1994; **35** (4): 316-327.
- 30. Williams JB, Gibbon M, First MB *et al.* The Structured Clinical Interview for DSM-III-R (SCID): II. Multisite Test-Retest Reliability. *Arch Gen Psychiat* 1992; **49** (8): 630-636.
- 31. Cochrane Effective Practice and Organisation of Care Review Group. Data Collection Checklist. Accessed 21/1/2011 from http://epoc.cochrane.org/epoc-resources-reviewauthors.
- 32. Costantini M, Musso M, Viterbori P *et al.* Detecting psychological distress in cancer patients: validity of the Italian version of the Hospital Anxiety and Depression Scale. *Support Care Cancer* 1999; **7** (3): 121-127.
- 33. Morasso G, Costantini M, Viterbori P *et al*. Predicting mood disorders in breast cancer patients. *Eur J Cancer* 2001; **37** (2): 216-223.
- 34. Walker J, Postma K, McHugh GS *et al.* Performance of the Hospital Anxiety and Depression Scale as a screening tool for major depressive disorder in cancer patients. *J Psychometric Res* 2007; **63**: 83-91.
- 35. Akechi T, Okuyama T, Sugawara Y *et al.* Screening for depression in terminally ill cancer patients in Japan. *J Pain Symptom Manag* 2006; **31** (1): 5-12.
- 36. Özalp E, Soygür H, Cankurtaran E *et al.* Psychiatric morbidity and its screening in Turkish women with breast cancer: a comparison between the HADS and SCID tests. *Psycho-Oncology* 2008; **17** (7): 668-675.
- 37. Singer S, Danker H, Dietz A *et al.* Screening for mental disorders in laryngeal cancer patients: a comparison of 6 methods. *Psycho-Oncology* 2008; **17** (3): 280-286.
- 38. Keller M, Sommerfeldt S, Fischer C *et al.* Recognition of distress and psychiatric morbidity in cancer patients: a multi-method approach. *Ann Oncol* 2004; **15**: 1243-1249.
- 39. Kugaya A, Akechi T, Okuyama T *et al.* Screening for Psychological Distress in Japanese Cancer Patients. *Jpn J Clin Oncol* 1998; **28** (5): 333-338.
- 40. Ravazi D, Delvaus N, Farvacques C *et al.* Screening for adjustment disorders and major depressive disorders in cancer in-patients. *Brit J Psychiat* 1990; **156**: 79-83.
- 41. Brown J,Buxton M. The economic perspective. *Br Med Bull* 1998; **54** (4): 993-1009.
- 42. Pignone MP, Gaynes BN, Rushton JL *et al.* Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002; **136**: 765-776.
- 43. AIHW (Australian Institute of Health and Welfare) & AACR (Australasian Association of Cancer Registries). Cancer in Australia: an overview. 2008, AIHW: Canberra
- 44. Surveillance Epidemiology and End Result (SEER). *SEER Stat Fact Sheets: Breast*. 2010; Available from: http://seer.cancer.gov/statfacts/html/breast.html.
- 45. Chen M-L,Tseng H-C. Symptom clusters in cancer patients. *Support Care Cancer* 2006; **14** (8): 825-830.

- 46. Coates A, Abraham S, Kaye SB *et al.* On the receiving end--patient perception of the sideeffects of cancer chemotherapy. *Eur J Cancer Clin On* 1983; **19** (2): 203-208.
- 47. Kirkova J,Walsh D. Cancer symptom clusters—a dynamic construct. *Support Care Cancer* 2007; **15** (9): 1011-1013.
- 48. Stark D, Kiely M, Smith A *et al.* Anxiety Disorders in Cancer Patients: Their Nature, Associations, and Relation to Quality of Life. *J Clin Oncol* 2002; **20** (14): 3137-3148.
- 49. Stark DPH, House A. Anxiety in cancer patients. Br J Cancer 2000; 83 (10): 1261-7.
- 50. Johnston M. Mood in chronic disease: Questioning the answers. *Current Psychol* 1999; **18** (1): 71-87.
- Mykletun A, Stordal E, Dahl AA. Hospital Anxiety and Depression (HAD) scale: factor structure, item analyses and internal consistency in a large population. *Brit J Psychiat* 2001; 179 (6): 540-544.
- 52. Kirmayer LJ. Cultural variations in the response to psychiatric disorders and emotional distress. *Soc Sci Med* 1989; **29** (3): 327-39.
- 53. Kirmayer LJ, Young A, Hayton BC. The cultural context of anxiety disorders. *Psychiatr Clin North Am* 1995; **18** (3): 503-21.
- 54. Lauber C,Rossler W. Stigma towards people with mental illness in developing countries in Asia. *Int Rev Psychiatr* 2007; **19** (2): 157-178.
- 55. Clarke DM,McKenzie DP. A caution on the use of cut-points applied to screening instruments or diagnostic criteria. *J Psychiat Res* 1994; **28** (2): 185-188.
- 56. Salander P. Motives that cancer patients in oncological care have for consulting a psychologist—an empirical study. *Psycho-Oncol* 2010; **19** (3): 248-54.
- 57. Sellick SM,Edwardson AD. Screenign new cancer patients for psychological distress using the hospital anxiety and depression scale. *Psycho-Oncol* 2007; **16**: 534-42.
- Curry C, Cossich T, Matthews J *et al*. Uptake of psychosocial referrals in an outpatient cancer setting: improving service accessibility via the referral process. *Support Care Cancer* 2002; **10** (7): 549-555.
- 59. Breen SJ, Baravelli CM, Schofield PE *et al.* Is symptom burden a predictor of anxiety and depression in patients with cancer about to commence chemotherapy? *Med J Australia* 2009; **190** (7 Suppl): S99-104.
- 60. DiMatteo MR, Lepper HS, Croghan TW. Depression is a risk factor for noncompliance with medical treatment. *Arch Intern Med* 2000; **160**: 2101-2107.
- 61. Katon W. Clinical and health services relationships between major depression, depressive symptoms, and general medical illness. *Biol Psychiat* 2003; **54**: 216-226.
- 62. Lepore SJ,Coyne JC. Psychological Interventions for Distress in Cancer Patients: A Review of Reviews. *Ann Behav Med* 2006; **32** (2): 85-92.
- 63. Jacobsen PB, Jim HS. Psychosocial Interventions for Anxiety and Depression in Adult Cancer Patients: Achievements and Challenges. *CA Cancer J Clin* 2008; **58** (4): 214-230.
- Kugaya A, Akechi T, Okuyama T *et al.* Prevalence, predictive factors, and screening for psychologic distress in patients with newly diagnosed head and neck cancer. *Cancer* 2000; 88 (12): 2817-2823.
- 65. Nelson CJ, Weinberger MI, Balk E *et al*. The chronology of distress, anxiety, and depression in older prostate cancer patients. *Oncologist* 2009; **14** (9): 891-9.
- Bauwens S, Baillon C, Distelmans W *et al.* The 'Distress Barometer': validation of method of combining the Distress Thermometer with a rated complaint scale. *Psycho-Oncol* 2009; **18** (5): 534-42.
- Boyes AW, Girgis A, Zucca AC *et al.* Anxiety and depression among long-term survivors of cancer in Australia: results of a population-based survey. *Med J Australia* 2009; **190** (7 Suppl): S94-8.
- 68. Caruso A, Vigna C, Marozzo B *et al.* Subjective versus objective risk in genetic counseling for hereditary breast and/or ovarian cancers. *J Exp Clin Canc Res* 2009; **28**: 157.

- Chen AM, Jennelle RLS, Grady V *et al.* Prospective study of psychosocial distress among patients undergoing radiotherapy for head and neck cancer. *Int J Radiat Oncol* 2009; **73** (1): 187-93.
- Delgado-Guay M, Parsons HA, Li Z *et al.* Symptom distress in advanced cancer patients with anxiety and depression in the palliative care setting. *Support Care Cancer* 2009; **17** (5): 573-9.
- 71. Mochizuki Y, Matsushima E, Omura K. Perioperative assessment of psychological state and quality of life of head and neck cancer patients undergoing surgery. *Int J Oral Max Surg* 2009; **38** (2): 151-9.
- Moorey S, Cort E, Kapari M *et al.* A cluster randomized controlled trial of cognitive behaviour therapy for common mental disorders in patients with advanced cancer. *Psychol Med* 2009; 39 (5): 713-23.
- 73. Noguera A, Centeno C, Carvajal A *et al*. Spanish "fine tuning" of language to describe depression and anxiety. *J Palliative Med* 2009; **12** (8): 707-12.
- 74. Potter S, Thomson HJ, Greenwood RJ *et al.* Health-related quality of life assessment after breast reconstruction. *Brit J Surg* 2009; **96** (6): 613-20.
- 75. Price MA, Zachariae R, Butow PN *et al*. Prevalence and predictors of insomnia in women with invasive ovarian cancer: anxiety a major factor. *Eur J Cancer* 2009; **45** (18): 3262-70.
- 76. Thomas BC, Thomas I, Nandamohan V *et al.* Screening for distress can predict loss of followup and treatment in cancer patients: results of development and validation of the Distress Inventory for Cancer Version 2. *Psycho-Oncology* 2009; **18** (5): 524-33.
- 77. van den Beuken-van Everdingen MH, de Rijke JM, Kessels AG *et al.* Quality of life and nonpain symptoms in patients with cancer. *J Pain Symptom Manag* 2009; **38** (2): 216-33.
- 78. Okuyama T, Endo C, Seto T *et al.* Cancer patients' reluctance to discuss psychological distress with their physicians was not associated with underrecognition of depression by physicians: a preliminary study. *Palliative Support Care* 2009; **7** (2): 229-33.
- 79. Schlich-Bakker KJ, ten Kroode HFJ, Warlam Rodenhuis CC *et al.* Distress in couples approached for genetic counseling and BRCA1/2 testing during adjuvant radiotherapy. *Psycho-Oncology* 2009; **18** (9): 965-73.
- 80. Clover K, Carter GL, Mackinnon A *et al.* Is my patient suffering clinically significant emotional distress? Demonstration of a probabilities approach to evaluating algorithms for screening for distress. *Support Care Cancer* 2009; **17** (12): 1455-62.
- Clover K, Leigh Carter G, Adams C *et al.* Concurrent validity of the PSYCH-6, a very short scale for detecting anxiety and depression, among oncology outpatients. *Aust NZ J Psychiat* 2009;
 43 (7): 682-8.
- 82. Hinz A, Krauss O, Stolzenburg J-U *et al.* Anxiety and depression in patients with prostate cancer and other urogenital cancer: a longitudinal study. *Urol Oncol* 2009; **27** (4): 367-72.
- 83. Low J, Gessler S, Williams R *et al.* Screening for distress and depression in cancer patients: is ultrashort depression screening a valid measure in the UK? A prospective validation study. *J Pain Symptom Manag* 2009; **38** (2): 234-43.
- Nelson C, Jacobson CM, Weinberger MI *et al.* The role of spirituality in the relationship between religiosity and depression in prostate cancer patients. *Ann Behav Med* 2009; **38** (2): 105-14.
- 85. Mhaidat NM, Alzoubi KH, Al-Sweedan S *et al*. Prevalence of depression among cancer patients in Jordan: a national survey. *Support Care Cancer* 2009; **17** (11): 1403-7.
- 86. Hartl DM, Dauchy S, Escande C *et al*. Quality of life after free-flap tongue reconstruction. *J Laryngol Otol* 2009; **123** (5): 550-4.

Table 1. Sensitivity and specificity of optimal thresholds (95% CI) recommended by studies assessing the validity of the HADS against the SCID in cancer

populations

Reference	Setting	Sample	HADS threshold recommended for	HADS threshold recommended for
			any psychological disorder or	major depression only
			adjustment disorder (with or without	
			major depression)	
Akechi et al.	Hospital Palliative Care	N = 209	T≥13 ^ª	T≥17
(2006)[35]	Unit	Mean age 61yrs	Sensitivity=0.80 (0.66-0.89)	Sensitivity=0.71 (0.42-0.86)
	Japan	66% male	Specificity=0.67 (0.59-0.74)	Specificity=0.77 (0.71-0.82)
		Terminally ill inpatients		
		Lung, colon, head and neck, liver cancers	D≥7 ^ª	D≥9
		No current cancer treatment	Sensitivity=0.78 (0.63-0.87)	Sensitivity=0.86 (0.56-0.94)
			Specificity=0.58 (0.50-0.65)	Specificity: 0.69 (0.62-0.75)
Costantini et al.	Two cancer treatment	<i>N</i> = 132	T≥10 ^c	
(1999)[32]/	centres	Mean age 53yrs	Sensitivity: 0.84 (NR)	
Morasso et al.	Italy	100% female	Specificity: 0.79 (NR)	
(2001)[33]		Outpatients receiving chemotherapy		
		Breast cancer; cancer stage not specified		
		All patients had completed		
		chemotherapy within past 12 months		
Keller <i>et al</i> .	Department of Surgery	<i>N</i> = 189	T≥16 [°]	
(2004)[38]	Germany	Mean age 57.4yrs	Sensitivity: 0.86 (NR)	
		60% male	Specificity: 0.87 (NR)	
		Inpatients admitted to undergo surgery		
		Heterogeneous cancer types (e.g.		
		colorectal, gastric, liver) and stage		
		Prior to undergoing surgery for cancer		
Kugaya <i>et al</i> .	Cancer Centre Hospital	<i>N</i> = 128	T≥10 ^b	T≥19
(1998)[39]	Japan	Mean age 61.1yrs	Sensitivity: 0.92 (NR)	Sensitivity: 0.82 (NR)
		62% male	Specificity: 0.66 (NR)	Specificity: 0.96 (NR)
		Majority (94%) inpatients	_	
		Heterogeneous cancer types (e.g.	A≥7 °	A≥7
		advanced lung, head and neck,	Sensitivity: 0.75 (NR)	Sensitivity: 0.94 (NR)
		digestive); cancer stage not specified	Specificity: 0.88 (NR)	Specificity: 0.88 (NR)

		85% pre-treatment		
			D≥4 ^b	D≥10
			Sensitivity: 0.92 (NR)	Sensitivity: 0.82 (NR)
			Specificity: 0.58 (NR)	Specificity: 0.95 (NR)
Kugaya <i>et al</i> .	Cancer Centre Hospital	<i>N</i> = 107	T≥15 ^b	
(2000)[64]	Japan	Mean age 61yrs	Sensitivity: 0.72 (NR)	
		76% male	Specificity: 0.81 (NR)	
		Inpatients		
		Head and neck cancers (oral cavity,		
		pharynx or larynx).		
		61% advanced cancer (stage III-IV)		
		All pre-treatment		
Ozalp et al.	Ankara Hospital	N = 204	T≥ 10 ^ª	T≥17
(2008)[36]	Turkey	Mean age 50.8yrs	Sensitivity: 0.84 (NR)	Sensitivity: 0.70 (NR)
		100% female	Specificity: 0.49 (NR)	Specificity: 0.80 (NR)
		Inpatients		
		Breast cancer	A≥5 ^ª	A≥7
		45% were pre-treatment; 34% had	Sensitivity: 0.88 (NR)	Sensitivity: 0.70 (NR)
		undergone surgery, 4% chemotherapy, 14% combined treatment	Specificity: 0.53 (NR)	Specificity: 0.65 (NR)
			D≥6 ^ª	D≥5
			Sensitivity: 0.72 (NR)	Sensitivity: 0.88 (NR)
			Specificity: 0.67 (NR)	Specificity: 0.60 (NR)
Ravazi et al.	Internal Medicine	N = 210	T≥13 ^b	T≥19
(1990)[40]	Department	Mean age 55.3yrs	Sensitivity: 0.75 (NR)	Sensitivity: 0.70 (NR)
	Belgium	67% female	Specificity: 0.75 (NR)	Specificity: 0.75 (NR)
		Inpatients		
		Heterogeneous cancer types and stage	A≥8 ^b	A≥11
		Treatment stages not specified	Sensitivity: 0.64 (NR)	Sensitivity: 0.54 (NR)
			Specificity: 0.72 (NR)	Specificity: 0.75 (NR)
			D≥7 ^b	D≥9
			Sensitivity: 0.59 (NR)	Sensitivity: 0.71 (NR)
			Specificity: 0.78 (NR)	Specificity: 0.76 (NR)
Singer et al.	Patient records from	N = 250	T≥14 ^c	T≥17
(2008)[37]	Leipzig tumour registry	Median age 60-69yrs	Sensitivity: 0.70 (NR)	Sensitivity: 0.85 (NR)

	Germany	91% male	Specificity: 0.80 (NR)	Specificity: 0.86 (NR)
		Outpatients		
		Laryngeal cancer	A≥7 [°]	A≥11
		All had undergone total or partial	Sensitivity: 0.72 (NR)	Sensitivity: 0.70 (NR)
		laryngectomy surgery in the past	Specificity: 0.80 (NR)	Specificity: 0.97 (NR)
			D≥ 7 [°]	D≥7
			Sensitivity: 0.67 (NR)	Sensitivity: 0.85 (NR)
			Specificity: 0.76 (NR)	Specificity: 0.73 (NR)
Singer et al.	Hospital clinics of	N = 689	T≥ 13 [°]	
(2009)[17]	University of Leipzig,	Median age 60-69	Sensitivity: 0.76 (NR)	
	Germany	59% male	Specificity: 0.60 (NR)	
	,	Inpatients		
		Heterogeneous cancer type and stage	A≥7 [°]	
		Heterogeneous treatment types	Sensitivity: 0.75 (NR)	
		с <i>п</i>	Specificity: 0.56 (NR)	
			D≥ 5 [°]	
			Sensitivity: 0.82 (NR)	
			Specificity: 0.49 (NR)	
Walker <i>et al</i> .	Outpatient clinics	N = 361		T≥ 15
(2007)[34]	UK	Mean age 61.7 vrs		Sensitivity: 0.87 (0.70-0.95)
(2007)[0.1]	•	77% female		Specificity: 0.85 (0.81-0.89)
		Outpatients		
		Heterogeneous cancer types and stages		۵>۹
		Treatment types not specified		Sensitivity: 0.87 (0.70-0.95)
		freutinent types not specified		Specificity: 0.83 (0.78-0.86)
				opeaniery: 0.05 (0.70 0.00)
				D≥7
				Sensitivity: 0.90 (0.74-0.97)
				Specificity: 0.88 (0.84-0.91)

NR Not Reported

a Adjustment disorder without major depression

b Adjustment disorder with major depression

c Any psychiatric disorder

Table 2. HADS thresholds for assessing 'psychological distress' used in the cancer prevalencestudies published in 2009 compared to the number of HADS cancer validation studies justifyingthe threshold used .

HADS scale	Threshold score used to	Proportion of validation studies
	determine prevalence	recommending the use of the same
	(proportion of prevalence	threshold score for 'any psychiatric
	studies) [reference]	disorder' or adjustment disorder
Anxiety subscale	A≥7 (1/24)[65]	3/10
	A≥8 (13/24)[59, 66-76]	1/10
	A≥11 (1/24)[77]	n/a (none)
Depression subscale	D≥7 (1/24)[65]	3/10
	D≥8 (13/24)[59, 66-76]	n/a (none)
	D≥11 (1/24)[77]	n/a (none)
Total score	T≥10 (1/24)[78]	3/10
	T≥13 (3/24)[17, 68, 79]	3/10
	T≥14 (1/24)[66]	n/a (none)
	T≥15 (4/24)[80-83]	1/10

N.B. Not all prevalence studies used both anxiety/depression subscale scores and the total score,

therefore not all numbers add to 24.

 Table 3. HADS thresholds for assessing depression used in the cancer prevalence studies published
 in 2009 compared to the threshold recommendations of HADS cancer validation studies.

HADS scale	Thresholds scores used to	Proportion of validation studies
	determine prevalence	recommending the use of the same
	(proportion of prevalence	threshold score for Major Depressive
	studies) [reference]	Disorder
Depression subscale	D≥7 (2/24)[83, 84]	2/10
	D≥8 (1/24)[85]	n/a (none)
	D≥10 (1/24)[86]	1/10
Total score	T≥19 (1/24)[68]	2/10
	T≥20 (1/24) [78]	n/a (none)