Using Rasch analysis to examine the Distress Thermometer's cut-off scores among a mixed group of patients with cancer

Sylvie D. Lambert RN, PhD
Ingram School of Nursing, McGill University, Australia

Julie F. Pallant PhD
Rural Health Academic Centre, University of Melbourne, Australia

Kerrie Clover PhD
Psycho-Oncology Service, Calvary Mater
Conjoint Lecturer, School of Psychology, Faculty of Science and Information Technology, University of Newcastle
Centre for Translational Neuroscience & Mental Health Research, University of Newcastle, Australia

Benjamin Britton DPsych
Psycho-Oncology Service, Calvary Mater Newcastle
Conjoint Lecturer, School of Medicine and Public Health, Faculty of Health, University of Newcastle
Centre for Translational Neuroscience & Mental Health Research, University of Newcastle, Australia

Madeleine T. King PhD
Cancer Australia's Chair in Cancer Quality of Life and
Director of the Quality of Life Office
Psycho-oncology Co-operative Research Group (PoCoG)
School of Psychology, The University of Sydney, Australia

Gregory Carter MBBS, FRANZCP, Cert Child Psych, PhD
A/Director, Department of Consultation-Liaison Psychiatry, Calvary Mater Newcastle Hospital. Conjoint Professor, Centre for Translational Neuroscience and Mental Health, School of Medical Practice and Population Health, Faculty of Health, University of Newcastle, Newcastle, New South Wales, Australia

Corresponding author:
Sylvie Lambert, Ingram School of Nursing, McGill University
Wilson Hall, 3506 Sherbrooke Street, Montreal, Canada
H3A 2A7
Tel: 514-967-3762
E-mail: sylvie.lambert@mcgill.ca

Abstract

Purpose: The Distress Thermometer (DT) is commonly used in cancer care to improve detection of distress. The DT's recommended cut-off score of 4 or 5 has typically been established using the Hospital Anxiety and Depression Scale (HADS) by Receiver Operating Characteristic curve analysis. The present analysis complements these studies by critically examining the use of the HADS to identify the DT's cut-off score and corroborating the DT's cut-off scores using Item Response Theory (Rasch analysis).

Methods: The DT and HADS were completed by 340 patients with cancer. Rasch dimensionality analysis was performed on the HADS-Total and test characteristic curves were examined to equate the DT and the HADS subscales. Identified DT cut-off scores were then examined for their sensitivity and specificity.

Results: Rasch analysis did not support the unidimensionality of HADS-Total. The test characteristic curves indicated that a cut-off score of $\geq$ 8 on the HADS-Anxiety and HADS-Depression subscales were equivalent to a score of 6 and 7 on the DT, respectively. However, a DT cut-off score of 5 resulted in the best balance between sensitivity and specificity across the HADS-Anxiety and HADS-Depression subscales.

Conclusions: Despite being a popular practice, the present findings did not support combining the HADS-Anxiety and HADS-Depression subscales to identify the DT's cut-off score. Furthermore, these results inform the use of the DT as a preliminary screening tool and suggests that when a single screen is used a DT cut-off score of 6 or 7 might be more appropriate than the typical cut-off score of 4.

Keywords: Psychometrics, anxiety, depression, cancer, psychological tests, questionnaire

Distress in cancer is typically defined as "an unpleasant emotional experience of a psychological (cognitive, behavioural, emotional), social, and/or spiritual nature that may interfere with the ability to cope effectively with cancer"[1], and is reported by 20% to 66% [2-5] of individuals with cancer. The negative consequences of elevated psychological distress are well-documented, and include difficulties in making decisions about treatment [1], lower treatment adherence [1; 6], lower quality of life [7], higher health risk behaviours [1], and higher severity and incidence of treatment side effects [6]. Although distress is acknowledged internationally as the sixth vital sign, it often goes unrecognized due to health care professionals' lack of time and confidence in assessing distress, which means that the majority of distressed patients are not referred to the specialists and services they need [1; 8].

Throughout the past decade, much attention has been given to implementing routine distress screening programs in cancer care as a way of improving the recognition and management of distress [1; 4; 9]. Although routine distress screening programs are generally acceptable to clinicians [10], a common concern is their feasibility in time and resource pressured oncology settings, including the appropriateness and practicality of the measure(s) used. It is well-recognised that ultra-short measures (less than three questions) are essential to ensure the acceptability of distress screening among clinicians [11]. The most popular ultra-short, screening measure of distress was introduced by The National Comprehensive Cancer Network's (NCCN's) Distress Management Guidelines - the Distress Thermometer (DT) [1]. The DT is quick and easy to use [12], provides a structured objective way to assess distress [13], allows for early identification of needs [13], and has been found to open up dialogue regarding issues that otherwise might not have been assessed [12; 14].

The DT is a visual analogue scale ranging from 0 = no distress to 10 = extreme distress. A Problem List is often used in conjunction with the DT to identify possible reasons contributing to the distress. To facilitate the identification of patients requiring additional

assessment and intervention, a number of studies have established the DT's cut-off score, relying on Receiver Operating Characteristic (ROC) curves analysis to compare its performance against once of the most widely used and validated measures of anxiety and depression, the Hospital Anxiety and Depression Scale (HADS) [15]. Using a cut-off score for borderline cases on the HADS-Anxiety and HADS-Depression subscales, a DT score of 4 or 5 is generally found to result in the best balance between sensitivity and specificity [16-19]. A number of additional studies have favoured a HADS-Total cut-off score of $\geq$ 15 to establish the DT's cut-off score of 4 or 5 [16; 17; 19-22]. This is despite confirmatory factor analyses finding that the HADS-Total typically fits less well than the original two-dimensional structure [23; 24]. Moreover, a recent Rasch analysis of the HADS-Total among cancer survivors did not support its unidimensionality [25], which suggests that the HADS-Anxiety and -depression subscales should not be combined to determine the DT's cut-off score.

Although ROC curve analysis is a popular and robust method for establishing the DT's cut-off score, this has been at the exclusion of other test equating methods. Of particular interest is the use of item response theory (IRT) or Rasch analysis (also referred to as the one parameter IRT model) to establish relationships between scores from two or more measures [26]. The strengths of Rach analysis are well-recognised and include sample invariance and that individuals are located on the same ability scale allowing comparisons between individuals and items [27]. Another advantage of using Rasch is that traditional analyses often assume interval-level data, which often is not the case. Rasch analysis does not assume item scores to be linear measures of the variable assessed [28].

The Rasch model posits that the probability of a person endorsing an item is a logistic function of the relative distance between the item location and the person's location on a linear scale. In other words, the probability that a person will endorse an item is a function of the difference between the person's level of, for example, distress and the 'level' of distress

expressed by the item. Rasch software expresses a person's level of ability on a 'logit' scale. Test equating with Rasch involves examining the test characteristic curves, in which individuals' scores from different tests are placed on the same ability (or logit) scale [26]. It is then possible to determine, for each level of, for example, distress the corresponding, equivalent score on each test. The process is analogous to equating kilograms to pounds by converting measurement from one scale to the other.

Given the central role of the DT in distress screening programs, and the reliance on cut-off scores to justify additional assessment and interventions, it is critical that other analytic methods, which complement ROC analyses, are used to strengthen conclusions about the cut-off scores used. Hence, using Rasch analysis, we critically examined the use of the HADS to confirm or update the DT's cut-off score previously determined with ROC curve studies. To our knowledge, this is the first study equating the DT and HADS using methods that extend beyond ROC curves.

Specific aims were to:

1) Despite the evidence discouraging the use of HADS-Total, it is commonly used to establish the cut-off score of the DT. Therefore, the first aim of this analysis was to determine the suitability of using the HADS-Total to identify the DT's cut-off score, using Rasch dimensionality analysis and test characteristic curves ;

2) Use test characteristic curves to identify DT cut-off scores that are equivalent to validated cut-off scores on the HADS; and

3) Calculate the accuracy statistics for various cut-off scores on the DT.

Materials and Methods

This study is a secondary analysis of data from a previous cross-sectional study to examine quality of life in oncology patients, which included the development of a two stage

algorithm to identify clinically significant emotional distress in cancer patients using the DT [29; 30]. The methods are briefly reviewed below, for additional information the reader is referred to Carter et al. [30] and Clover et al. [29].

Patients and setting

The study was conducted at the Calvary Mater Newcastle hospital, NSW, Australia during an 8-week period (April–May 2005). This hospital provides comprehensive secondary and tertiary treatment services for adults with malignant disease in Medical Oncology, Surgical Oncology, Radiation Oncology, and Haematology. Adult patients with sufficient English language skills, attending non-surgical Oncology and Haematology outpatient clinics, and well-enough to participate in this study were eligible. The main exclusion criterion was attending the clinic for the first time. The Hunter Area Health Service Human Research Ethics Committee approved the original study and this secondary analysis.

Procedures

Potentially eligible patients were given an information slip by the reception staff at the time of checking in for their appointment, and advised to keep the slip if they were interested in hearing more about the study. People holding the information slip were then approached in the waiting room by a trained hospital volunteer or a member of the research team and given more information about the study. Patients who were eligible and agreed to participate in the study provided written consent and were then invited to complete the study measures.

Data collection

While waiting for their clinical appointment, participants completed a computer-administered battery of questionnaires, which included the DT and HADS.

Distress thermometer. The DT is a visual analogue scale ranging from 0 = no distress to 10 = extreme distress [1]. Participants were instructed to select the number that best described the overall level of distress they experienced in the past week.

Hospital Anxiety and Depression Scale. The HADS is a 14-item self-administered questionnaire, with equal number of items measuring anxiety (HADS-Anxiety subscale) and depression (HADS-Depression subscale). Each item is rated on a four-point response scale (0 to 3 – variable responses), where a higher score indicates more anxiety/depression. Subscale scores are obtained by summing component items and these are typically categorized as normal (scores less than 8), borderline or possible cases (scores of 8 to 10), and clinical or probable cases (scores of 11 to 21) [15]. In addition to the HADS-Depression and HADS-Anxiety subscales, in psycho-oncology it has become increasingly popular to use the HADS-Total, as a unidimensional measure of emotional distress, with a cut-off score of $\geq$ 15 [16; 17; 19-21]. Although HADS-Total has gained some popularity [19], a number of confirmatory factor analyses have not supported its use [23; 24].

Clinical information was extracted from participants' medical records with their written consent by a qualified medical practitioner. Information collected included cancer type, stage, and current (at the time of interview) treatment with chemotherapy, radiotherapy and/or other.

Data Analysis

Data analysis was conducted in stages. Firstly, dimensionality was assessed, using the partial credit model of the Rasch Unidimensional Models for Measurement (RUMM) software version 2030 [31], to determine whether it is suitable to combine the two HADS subscales to

form the HADS-Total. To this end, the 14 items of the HADS were subjected to Principal Component Analysis (PCA) of the residuals to identify the two subsets of items that showed the greatest difference from one another. A series of t-tests were then used to assess the differences between person estimates (location values) derived from these two subsets of items. If less than 5% of these tests were significantly different (or specifically the lower bound of the binomial confidence interval was less than 5%), the scale was considered unidimensional (i.e., implies that only one construct is measured by a set of items in a scale) [28].

Secondly, test equating analysis was conducted, again using RUMM2030. This involved creating subtests corresponding to the DT and each HADS subscale and examining the test characteristic curves to determine the equivalent cut-off scores (also referred to as co-calibration). Test characteristic curves plot participants' raw scores on a scale (y axis) against their ability or location (x axis) on the measured trait for each subtest. RUMM2030 expresses participants' ability (i.e., distress, anxiety, depression) on a common logit scale. From this, it was possible to equate HADS-Anxiety or HADS-Depression subscales and the DT to both HADS subscales.

Finally, accuracy statistics for various cut-off scores on the DT were calculated [32], including the positive predictive value (PPV, the probability that the patient has the disease when restricted to those patients who test positive), the negative predictive value (NPV, the probability that a patient with a negative test result really is free of the condition for which the test was conducted), sensitivity (ability of a test to correctly identify those individuals with the condition), specificity (ability of a test to correctly identify those without the condition or those that are well), and accuracy (the percentage of test results correctly identified by the test) were

calculated using SPSS version 19 for the cut-off scores obtained from the test equating analysis.

Results

Participants

Of the 1707 patients who attended the clinics during the recruitment period, 393 participants consented to study participation; however, complete data were obtained for 377. Of these, 340 participants had malignant disease and were included in this analysis. Full details of the sample has been reported elsewhere [33]. Participants mean age was 60 years (standard deviation = 12, range = 18-88 years), and an almost equivalent proportion of men and women were included (52% male). Most participants (74%) were married or in a de facto relationship. Almost half of the sample was retired (49%). Two thirds of participants were receiving treatment at the time of the study, and the three most common cancer diagnoses were haematological (24%), breast (24%), or urological cancer (15%).

Participants' scores on the HADS and DT

The mean scores for the study measures are summarized in Table 1. Slightly more than a third of participants scored $\geq$ 8 on the HADS-Anxiety subscale and 22.1% scored $\geq$ 8 on the HADS-Depression. Almost half of participants reported a DT score $\geq$ 4.

INSERT TABLE 1 ABOUT HERE

Dimensionality of the HADS-Total

Principal component analysis of the residuals defined two subsets of HADS items that closely mirrored those of the original HADS-Anxiety and -Depression subscales, with the exception of item #7 "*I can sit at ease and feel relaxed*". A series of t-tests performed on the

person estimates from these two subsets of items revealed that 14.06% (95% CI = 11.86%-16.26%) of participants had significantly different t-tests. This indicates that the HADS-Total is not a unidimensional measure of distress. Hence, the calibration of the DT was done using the two HADS subscales separately. Furthermore, the test characteristic curves indicated that the cut-off of 8 on the HADS-Anxiety subscale equated to a score of 6 (not 8) on the HADS-Depression subscale (Figure 1).

INSERT FIGURE 1 ABOUT HERE

Equating the DT and HADS subscales

HADS-Anxiety and DT

Using the test characteristic curves (Figure 2), the 'borderline' cut-off score of 8 on the HADS-Anxiety is equivalent to 5.5 (or 6 rounded up) on the DT, and the 'clinical' cut-off score of 11 is equivalent to 7 on the DT. Table 2 lists the equivalent scores for all other HADS-Anxiety and DT scores.

INSERT FIGURE 2 AND TABLE 2 ABOUT HERE

HADS-Depression and DT

The test characteristic curves showed that the 'borderline' cut-off score of 8 on the HADS-Depression is equivalent to 6.5 (or 7 rounded up) on the DT (Figure 3), and the 'clinical' cut-off score of 11 is equivalent to 8 on the DT. Table 2 lists the equivalent scores for all other HADS-Depression and DT scores.

INSERT FIGURE 3 ABOUT HERE

Determining the best cut-off score for the DT

HADS-Anxiety and DT

As presented in Table 3, a DT score of ≥ 5 resulted in the best balance between sensitivity (74.4%) and specificity (76.6%) for detecting cases of anxiety. However, the above test equating analyses suggested that a DT score between 5 and 6 is equivalent to HADS-Anxiety ≥ 8. A DT cut-off score of 6 correctly identified 57.3% of HADS cases of anxiety (i.e., score ≥ 8) and 86.0% of non cases. At this cut-off score, 68.4% of the anxiety cases identified by the DT were cases according to the HADS. Of the non-cases identified by the DT, 79.3% were non-cases according to the HADS.

HADS-Depression and DT

For depression, a DT score of ≥ 5 also resulted in the best balance between sensitivity (72.0%) and specificity (67.5%) for detecting cases of depression. The above test-equating analyses suggested that a DT score between 6 and 7 is equivalent to HADS-Depression ≥ 8. A DT score of 6 identified 58.7% of HADS cases of 'depression', which decreased to 46.7% with a DT score of 7. A DT cut-off score of 6 identified slightly more than three quarters of HADS-Depression non-cases and a score of 7 identified 85.7% of non-cases.

INSERT TABLE 3 ABOUT HERE

Discussion

The aim of this paper was to calibrate one of the most well-known, ultra-short questionnaires in oncology – the DT - against the HADS, one of the most widely used instruments in oncology for detecting anxiety or depression [34]. To date, ROC curve analysis has been the most popular technique to establish the DT's cut-off score, and the current study

complements this literature by documenting equivalent scores across the DT and HADS using Rasch Analysis. Key findings from this analysis were 1) the HADS-Total is not a unidimensional scale and therefore its use in 'validating' other scales or as a score for detecting psychological distress is not supported and 2) a score of 6 (rounded) and 7 (rounded) on the DT are equivalent to the cut-off score of ≥ 8 on the HADS-Anxiety and HADS-Depression subscales, respectively.

Using the HADS-Total to validate the DT

In the psycho-oncology literature, all 14 HADS items are often considered together as a unidimensional measure of emotional distress and used to determine the DT's cut-off score [17; 19; 20]. The present Rasch analysis did not support the unidimensionality of the HADS-Total, corroborating findings from other Rasch analyses [25; 35] and confirmatory factor analyses [23; 24], and cautions against using the HADS-Total. This findings is also consistent with a commentary by Snaith [36] stressing that the HADS was developed to distinguish between the constructs of anxiety and depression and not as a screening tool for overall distress. A review by Bjelland et al.[37] found that the correlations between the two HADS subscales varied from 0.40 to 0.74 (mean 0.56), which recognises the comorbidity between anxiety and depression, but also that these phenomena have characteristics that are distinct from each other.

Appropriateness of the DT cut-off score of ≥ 4

All distress screening programs need to balance a trade-off between sensitivity and specificity against available resources. In the present study, at the cut-off score of 4, the DT detected 82.1% of individuals with anxiety. Where access to psychosocial service is readily available, it might be possible to use a DT cut-off score with high sensitivity so that cases of

distress are not missed. However, most settings are pressured for time and resources, and such an inclusive approach might inadvertently lead to clinician burden, overloading of psychosocial services, and adversely impacting on service access for those most in need. Although the DT is appropriate in *ruling out* depression, anxiety, or distress, it falls short when it comes to *ruling in* these conditions [38]. Thus, when using the DT as part of a distress screening program, our findings and those of others [38] suggest that either: 1) the DT is part of a two-stage screening process, if the DT cut-off score of ≥ 4 is used or 2) a higher cut-off score is used, if a two-stage screening process is not feasible.

Use the DT cut-off score of ≥ 4, if a two-stage screening process is feasible

Our findings corroborated those of others [39], and emphasised that a DT cut-off score of ≥ 4 led to over-detection of cases. The suggestion then is to use a two-stage screening process [39], where only individuals scoring above the threshold complete a second test with higher specificity. Such an approach acknowledges that the DT is a screening measure (not a diagnostic one) and its main purpose is to delineate those who might need to be further assessed. The advantages of a two-stage approach include maximising accuracy and minimising the number of items that all patients need to complete, which in turn is more time efficient for clinicians [29; 40] and less stressful to patients. Clover et al. [29] with the clinical population included in the current study, demonstrated that administering the PSYCH-6 only to those patients scoring ≥ 4 on the DT was superior to the DT alone and equally accurate in case identification as the HADS, but practically superior (fewer items and lower cost).

Use a higher cut-off score, if a two-stage screening process is not feasible

In many setting, a two-stage screening program might not be feasible. Mitchell et al. [11] found that clinicians prefer to use one very short instrument (three questions or less) to

screen for all psychiatric conditions, or no method at all. In settings where resources are limited, using a cut-off score with higher specificity might be necessary to direct resources towards patients likely to be experiencing high levels of distress. This approach will inadvertently miss cases of distress; however, if available services are already overloaded, it is unproductive and perhaps unethical to identify cases of distress that cannot be assisted. Our results suggest that in these settings a DT cut off score of 6 (for anxiety) or 7 (for depression) might be appropriate.

Another suggestion to deal with the various cut-off scores of the DT is to use a stepped care approach [41]. For instance, Hoffman et al., [41] based on ROC curve analyses, found that no one score on the DT maximised both sensitivity and specificity, and suggested that when DT scores fall between 0 and 3, no additional action is needed, whereas scores between 4 and 6 could trigger referral to less intensive sources of psychosocial assistance (e.g., self-help group), and scores $\geq$ 7 could result in referral to more intensive interventions such as an individual appointment with a psychologist. A stepped care approach in cancer is especially relevant, as most patients report mild to moderate distress, and a number of minimal interventions (e.g., self-administered stress management interventions) have been found to be efficacious in alleviating distress [42-45] and cost-effective [44]. The recommendation that a score of $\geq$ 7 could result in a referral to psychosocial services is consistent with the findings from this study.

Strengths and weaknesses

The strengths of the study are its adequate sample size for Rasch analysis [46], involvement of a mixed group of oncology outpatients, including an almost equivalent proportion of men and women, and use of modern Rasch analysis to establish equivalent scores across the HADS subscales and the DT. One limitation of this analysis is reliance on the HADS to co-calibrate the DT. Although the HADS is the most widely used scale to calibrate the DT, a

number of studies [34; 37; 47-49], including Rasch analyses [25], have reported on its psychometric weaknesses, including a lack of consensus on the optimal cut-point for cases of anxiety and depression and problematic fit of some items. Another limitation is the low response rate (23%). Some of the statistics used to evaluate the appropriateness of the DT cut-off scores are influenced by the prevalence of the condition in the sample, thus the findings of this study might not necessarily apply to other samples.

Implications and conclusion

Findings from the present study did not support the use of the HADS-Total to calibrate the DT. Furthermore, the recommended DT cut-off score of ≥4 was not supported, and higher DT scores were found to be equivalent to the borderline cut-off score on the HADS-Anxiety (DT cut-off score of 6) and HADS–Depression (DT cut-off score of 7). Clinically, these results emphasised the role of the DT as a preliminary screening tool, where a DT score of ≥ 4 score might be appropriate when a two-stage screening process is feasible and increasing the cut-off score to 6 or 7 might be most appropriate for time- and resource-pressured environments. Future research is needed to corroborate our calibration findings using Rasch analysis and explore the impact of different distress care pathways on care and outcomes.

Table captions

Table 1 *Description of study measures*

Table 2 *Equivalent scores on the DT, HADS-Anxiety, and HADS-Depression*

Table 3 *Accuracy statistics for DT cut-off scores versus HADS-Anxiety and HADS-Depression subscales*

Figure captions

Fig 1 Equating of scores for the HADS-Anxiety (line 1) and HADS-Depression (line 2). Vertical or y axis = score of the HADS-Anxiety or HADS-Depression and horizontal or x axis = location on the logit scale or level of ability. Read a score of 8 across to the HADS-Anxiety (line 1); then read down to HADS-Depression (line 2), and then read back to the raw score, which is 6

Fig 2 Equating of scores for the HADS-Anxiety (line 1) and DT (line 2). Vertical or y axis = score on the HADS-Anxiety or DT and horizontal or x axis = location on the logit scale or level of ability. Read a score of 8 across to the HADS-Anxiety (line 1); then read down to DT (line 2), and then read back to the raw score, which is 6 (rounded)

Fig *3* Equating of scores for the HADS-Depression (line 1) and DT (line 2). Vertical or y axis = score on the HADS-Depression or DT and horizontal or x axis = location on the logit scale or level of ability. Read a score of 8 across to the HADS-Depression (line 1); then read down to DT (line 2), and then read back to the raw score, which is 6.5 or 7 rounded

Acknowledgements

Funding sources: Dr Lambert was supported by a NHMRC Research Fellowship (APP1012869).

References

1. National Comprehensive Cancer Network. (2012). Distress management clinical practice guidelines (Vol. 2.2013): National Comprehensive Cancer Network.
2. National Breast Cancer Centre, National Cancer Control Initiative. (2003). Clinical Practice Guidelines for the Psychosocial Care of Adults with Cancer. Camperdown, NSW: National Breast Cancer Centre.
3. Carlson, L. E., Angen, M., Cullum, J., Goodey, E., Koopmans, J., Lamont, L., MacRae, J. H., Martin, M., Pelletier, G., Robinson, J., Simpson, J. S. A., Speca, M., Tillotson, L., & Bultz, B. D. (2004). High levels of untreated distress and fatigue in cancer patients. British Journal of Cancer, 90, 2297-2304.
4. Zabora, J., BrintzenhofeSzoc, K., Curbow, B., Hooker, C., & Plantadosi, S. (2001). The prevalence of psychological distress by cancer site. Psycho-Oncology, 10, 19-28.
5. Mitchell, A. J., Meader, N., & Symonds, P. (2010). Diagnostic validity of the Hospital Anxiety and Depression Scale (HADS) in cancer and palliative settings: A meta-analysis. Journal of Affective Disorders, 126(3), 335-348.
6. Rabin, E. G., Heldt, E., Hirakata, V. N., Bittelbrunn, A. C., Chachamovich, E., & Fleck, M. P. A. (2009). Depression and perceptions of quality of life of breast cancer survivors and their male partners. Oncology Nursing Forum, 36(3), E153-158.
7. Head, B. A., Schapmire, T. J., Keeney, C. E., Deck, S. M., Studts, J. L., Hermann, C. P., Scharfenberger, J. A., & Pfeifer, M. P. (2012). Use of the Distress Thermometer to discern clinically relevant quality of life differences in women with breast cancer. Quality of Life Research, 21(2), 215-223.
8. Holland, J. C. (2004). IPOS Sutherland Memorial Lecture: an international perspective on the development of psychosocial oncology: overcoming cultural and attitudinal barriers to improve psychosocial care. Psycho-Oncology, 13, 445-459.
9. Carlson, L. E., & Bultz, B. D. (2003). Cancer distress screening - needs, models, and methods. Journal of Psychosomatic Research, 55(5), 403-409.
10. Mitchell, A. J., Lord, K., Slattery, J., Grainger, L., & Symonds, P. (2012). How feasible is implementation of distress screening by cancer clinicians in routine clinical care? Cancer, 118(24), 6260-6269.
11. Mitchell, A. J., Kaar, S., Coggan, C., & Herdman, J. (2008). Acceptability of common screening methods used to detect distress and related mood disorders—preferences of cancer specialists and non-specialists. Psycho-Oncology, 17(3), 226-236.
12. Fulcher, C. D., & Gosselin-Acomb, T. K. (2007). Distress assessment: practice change through guideline implementation. Clin J Oncol Nurs, 11(6), 817-821.
13. Hughes, K., Sargeant, H., & Hawkes, A. (2011). Acceptability of the Distress Thermometer and Problem List to community-based telephone cancer helpline operators, and to cancer patients and carers. BMC Cancer, 11(1), 1-8.
14. Lambert, S. D., Kelly, B., Boyes, A., Cameron, A., Adams, C., Proietto, A., & Girgis, A. (Accepted). Insights into preferences for psycho-oncology services among women with gynaecologic cancer following distress screening. Journal of the National Comprehensive Cancer Network.
15. Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. Acta Psychiatrica Scandinavica, 67(6), 361-370.
16. Baken, D. M., & Woolley, C. (2011). Validation of the Distress Thermometer, Impact Thermometer and combinations of these in screening for distress. Psycho-Oncology, 20(6), 609-614.
17. Gessler, S., Low, J., Daniells, E., R, W., Brough, V., Tookman, A., & Jones, L. (2008). Screening for distress in cancer patients: is the distress thermometer a valid measure in

the UK and does it measure change over time? A prospective validation study. Psycho-Oncology, 17, 538-547.
18. Low, J., Gessler, S., Williams, R., Daniells, E., Brough, V., Tookman, A., & Jones, L. (2009). Screening for Distress and Depression inÂ Cancer Patients: Is Ultrashort Depression Screening a Valid Measure in the UK? A Prospective Validation Study. Journal of Pain and Symptom Management, 38(2), 234-243.
19. Jacobsen, P. B., Donovan, K. A., Trask, P. C., Fleishman, S. B., Zabora, J., Baker, F., & Holland, J. C. (2005). Screening for psychologic distress in ambulatory cancer patients: a multicenter evaluation of the distress thermometer. Cancer, 103, 1494-1502.
20. Gil, F., Grassi, L., Travado, L., Tomamichel, M., & Gonzalez, J. R. (2005). Use of distress and depression thermometers to measure psychosocial morbidity among southern European cancer patients. Support Care Cancer, 13, 600-606.
21. Ozalp, E., Cankurtaran, E. S., Soygur, H., Geyik, P. O., & Jacobsen, P. B. (2007). Screening for psychological distress in Turkish cancer patients. Psychooncology, 16(4), 304-311.
22. Grassi, L., Johansen, C., Annunziata, M. A., Capovilla, E., Costantini, A., Gritti, P., Torta, R., & Bellani, M. (2013). Screening for distress in cancer patients: A multicenter, nationwide study in Italy. Cancer, 119(9), 1714-1721.
23. Dunbar, M., Ford, G., Hunt, K., & Der, G. (2000). A confirmatory factor analysis of the Hospital Anxiety and Depression Scale: Comparing empirically and theoretically derived structures. British Journal of Clinical Psychology, 39, 73-94.
24. Rodgers, J., Martin, C., Morse, R., Kendell, K., & Verrill, M. (2005). An investigation into the psychometric properties of the Hospital Anxiety and Depression Scale in patients with breast cancer. Health and Quality of Life Outcomes, 3(1), 41.
25. Lambert, S., Pallant, J. F., Boyes, A., King, M., Britton, B., & Girgis, A. (2013). A Rasch analysis of the Hospital Anxiety and Depression Scale (HADS) among cancer survivors. Psychological Assessment, 25(2), 379-390.
26. Zhu, W. (1998). Test equating: what, why, how? Research Quarterly for Exercise & Sport, 69(1), 11-23.
27. Bortolotti, S., Tezza, R., Andrade, D., Bornia, A., & Sousa, J. A. (2013). Relevance and advantages of using the item response theory. Quality & Quantity, 47(4), 2341-2360.
28. Tennant, A., & Conaghan, P. G. (2007). The Rasch Measurement Model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis and Rheumatism, 57(8), 1358-1362.
29. Clover, K., Carter, G., Mackinnon, A., & Adams, C. (2009). Is my patient suffering clinically significant emotional distress? Demonstration of a probabilities approach to evaluating algorithms for screening for distress. Supportive Care in Cancer, 17(12), 1455-1462.
30. Carter, G., Lewin, T., Rashid, G., Adams, C., & Clover, K. (2008). Computerised assessment of quality of life in oncology patients and carers. Psycho-Oncology, 17, 26-33.
31. Andrich, D., Lyne, A., Sheridan, B., & Luo, G. (2010). RUMM2030. Perth, Western Australia.
32. Loong, T. W. (2003). Understanding sensitivity and specificity with the right side of the brain. BMJ, 327(7417), 716-719.
33. Clover, K., Leigh Carter, G., Adams, C., Hickie, I., & Davenport, T. (2009). Concurrent validity of the PSYCH-6, a very short scale for detecting anxiety and depression, among oncology outpatients. Australian and New Zealand Journal of Psychiatry, 43(7), 682-688.

34. Luckett, T., Butow, P. N., King, M. T., Oguchi, M., Heading, G., Hackl, N. A., Rankin, N., & Price, M. A. (2010). 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. Supportive Care Cancer, 18, 1241-1262.
35. Lambert, S., Pallant, J. F., & Girgis, A. (2011). Rasch analysis of the Hospital Anxiety and Depression Scale among caregivers of cancer survivors: implications for its use in psycho-oncology. Psycho-Oncology, 20(9), 919-925.
36. Snaith, R. (1991). The GHQ and the HAD. The British Journal of Psychiatry, 158(3), 433a-.
37. Bjelland, I., Dahl, A., Haug, T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale: an updated literature review. Journal of Psychosomatic Research, 52(2), 69-77.
38. Mitchell, A. J. (2007). Pooled results from 38 analyses of the accuracy of distress thermometer and other ultra-short methods of detecting cancer-related mood disorders. Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 25(29), 4670-4681.
39. Boyes, A., D'Este, C., Carey, M., Lecathelinais, C., & Girgis, A. (2012). How does the Distress Thermometer compare to the Hospital Anxiety and Depression Scale for detecting possible cases of psychological morbidity among cancer survivors? Supportive Care in Cancer, 1-9.
40. Bagha, S. M., Macedo, A., Jacks, L. M., Lo, C., Zimmermann, C., Rodin, G., & Li, M. (2013). The utility of the Edmonton Symptom Assessment System in screening for anxiety and depression. European Journal of Cancer Care (Engl), 22(1), 60-69.
41. Hoffman, B. M., Zevon, M. A., D'Arrigo, M. C., & Cecchini, T. B. (2004). Screening for distress in cancer patients: the NCCN rapid-screening measure. Psycho-Oncology, 13(11), 792-799.
42. Beatty, L., Koczwara, B., Rice, J., & Wade, T. (2010). A randomised controlled trial to evaluate the effects of a self-help workbook intervention on distress, coping and QOL after breast cancer diagnosis. Medical Journal of Australia, 193(5 Suppl), S68-73.
43. Beatty, L., Koczwara, B., & Wade, T. (2011). 'Cancer Coping Online': A pilot trial of a self-guided CBT internet intervention for cancer-related distress. E-Journal of Applied Psychology, 7(2), 17 - 25.
44. Jacobsen, P. B., Meade, C. D., Stein, K. D., Chirikos, T. N., Small, B. J., & Ruckdeschel, J. C. (2002). Efficacy and costs of two forms of stress management training for cancer patients undergoing chemotherapy. Journal of Clinical Oncology, 20(12), 2851-2862.
45. Jacobsen, P. B., Phillips, K. M., Jim, H. S. L., Small, B. J., Faul, L. A., Meade, C. D., Thompson, L., Williams, C. C., Loftus, L. S., Fishman, M., & Wilson, R. W. (2013). Effects of self-directed stress management training and home-based exercise on quality of life in cancer patients receiving chemotherapy: a randomized controlled trial. Psycho-Oncology, 22(6):1229-1235.
46. Chen, W.-H., Lenderking, W., Jin, Y., Wyrwich, K., Gelhorn, H., & Revicki, D. (2013). Is Rasch model analysis applicable in small sample size pilot studies for assessing item characteristics? An example using PROMIS pain behavior item bank data. Quality of Life Research, 1-9.
47. Walker, J., Postma, K., McHugh, G. S., Rush, R., Coyle, B., Strong, V., & Sharpe, M. (2007). Performance of the Hospital Anxiety and Depression Scale as a screening tool for major depressive disorder in cancer patients. Journal of Psychosomatic Research, 63(1), 83-91.

48. Singer, S., Kuhnt, S., Götze, H., Hauss, J., Hinz, A., Liebmann, A., Krauss, O., Lehmann, A., & Schwarz, R. (2009). Hospital anxiety and depression scale cutoff scores for cancer patients in acute care. British Journal of Cancer, 100(6), 908-912.
49. Vodermaier, A., & Roanne, D. M. (2011). Accuracy of the Hospital Anxiety and Depression Scale as a screening tool in cancer patients: A systematic review and meta-analysis. Supportive Care in Cancer, 19(12), 1899-1908.

Using Rasch analysis to examine the Distress Thermometer's cut-off scores among a mixed group of patients with cancer

Abstract

Purpose: The Distress Thermometer (DT) is commonly used in cancer care to improve detection of distress. The DT's recommended cut-off score of 4 or 5 has typically been established using the Hospital Anxiety and Depression Scale (HADS) by Receiver Operating Characteristic curve analysis. The present analysis complements these studies by critically examining the use of the HADS to identify the DT's cut-off score and corroborating the DT's cut-off scores using Item Response Theory (Rasch analysis).

Methods: The DT and HADS were completed by 340 patients with cancer. Rasch dimensionality analysis was performed on the HADS-Total and test characteristic curves were examined to equate the DT and the HADS subscales. Identified DT cut-off scores were then examined for their sensitivity and specificity.

Results: Rasch analysis did not support the unidimensionality of HADS-Total. The test characteristic curves indicated that a cut-off score of $\geq$ 8 on the HADS-Anxiety and HADS-Depression subscales were equivalent to a score of 6 and 7 on the DT, respectively. However, a DT cut-off score of 5 resulted in the best balance between sensitivity and specificity across the HADS-Anxiety and HADS-Depression subscales.

Conclusions: Despite being a popular practice, the present findings did not support combining the HADS-Anxiety and HADS-Depression subscales to identify the DT's cut-off score. Furthermore, these results inform the use of the DT as a preliminary screening tool and suggests that when a single screen is used a DT cut-off score of 6 or 7 might be more appropriate than the typical cut-off score of 4.

Keywords: Psychometrics, anxiety, depression, cancer, psychological tests, questionnaire

Table 1. Description of study measures

Measure	Mean score (SD)	No. of patients (%)	Score Range	
			Possible	Observed
DT	3.71 (2.93)	--	0-10	0-10
scoring ≥ 4	--	161 (47.4)		
HADS				
HADS-Anxiety	6.27 (4.27)	--	0-21	0-21
scoring ≥ 8	--	117 (34.5)		
scoring ≥ 11	--	64 (18.9)		
HADS-Depression	4.66 (3.64)	--	0-21	0-16
scoring ≥ 8	--	75 (22.1)		
scoring ≥ 11	--	35 (10.3)		

Table 2. Equivalent scores on the DT, HADS-Anxiety, and HADS-Depression

DT	HADS-Anxiety	HADS-Depression
1	3	1.5
2	4	2.5
3	5	3
4	6	4
5	7	5.5
6	9	7
7	11	9
8	13	11
9	15.5	12.5
10	17.5	14

Table 3. Accuracy statistics for DT cut-off scores versus HADS-Anxiety and HADS-Depression subscales

	Positive Predictive Value (95% CI)	**Negative Predictive Value (95% CI)**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**Accuracy (95% CI)**
HADS-Anxiety and DT					
DT ≥ 3	54.1 (47.1 – 61.1)	95.5 (90.1 – 98.2)	94.9 (88.7 – 97.9)	57.7 (50.9 – 64.2)	70.5 (65.3 – 75.2)
DT ≥ 4	60.0 (51.9 – 67.6)	88.3 (82.4 – 92.4)	82.1 (73.6 – 88.3)	71.2 (64.7 – 76.9)	74.9 (69.9 – 79.4)
DT ≥ 5	62.6 (53.9 – 70.5)	85.0 (79.1 – 89.5)	74.4 (65.3 – 81.8)	76.6 (70.3 – 81.9)	75.8 (70.8 – 80.2)
DT ≥ 6	68.4 (58.1 – 77.2)	79.3 (73.5 – 84.1)	57.3 (47.8 – 66.3)	86.0 (80.6 – 90.2)	76.1 (71.1 – 80.5)
DT ≥ 7	75.0 (63.1 – 84.1)	76.4 (70.8 – 81.3)	46.2 (37.0 – 55.6)	91.9 (87.2 – 95.0)	76.1 (71.1 – 80.5)
DT ≥ 8	75.6 (59.4 – 87.1)	71.1 (65.6 – 76.1)	26.5 (19.0 – 35.6)	95.5 (91.6 – 97.7)	71.7 (66.5 – 76.4)
HADS-Depression and DT					
DT ≥ 3	33.5 (27.2 – 40.4)	95.5 (90.1 – 98.2)	92.0 (82.8 – 96.7)	48.3 (42.2 – 54.5)	57.9 (52.5 – 63.2)
DT ≥ 4	37.3 (29.9 – 45.3)	91.6 (86.3 – 95.1)	80.0 (68.9 – 88.0)	61.9 (55.7 – 67.7)	65.9 (60.5 – 70.9)
DT ≥ 5	38.6 (30.6 – 47.2)	89.5 (84.2 – 93.2)	72.0 (60.3 – 81.5)	67.5 (61.5 – 73.1)	68.5 (63.3 – 73.4)
DT ≥ 6	44.4 (34.6 – 54.8)	87.1 (82.0 – 91.0)	58.7 (46.7 – 69.7)	79.2 (73.8 – 83.9)	74.7 (69.7 – 79.2)
DT ≥ 7	47.9 (36.2 – 59.9)	85.0 (80.0 – 89.0)	46.7 (35.2 – 58.5)	85.7 (80.7 – 89.5)	77.1 (72.2 – 81.4)
DT ≥ 8	48.8 (33.2 – 64.6)	81.6 (76.6 – 85.7)	26.7 (17.4 – 38.3)	92.1 (88.0 – 94.9)	77.7 (72.8 – 81.9)

Note. CI – confidence interval.

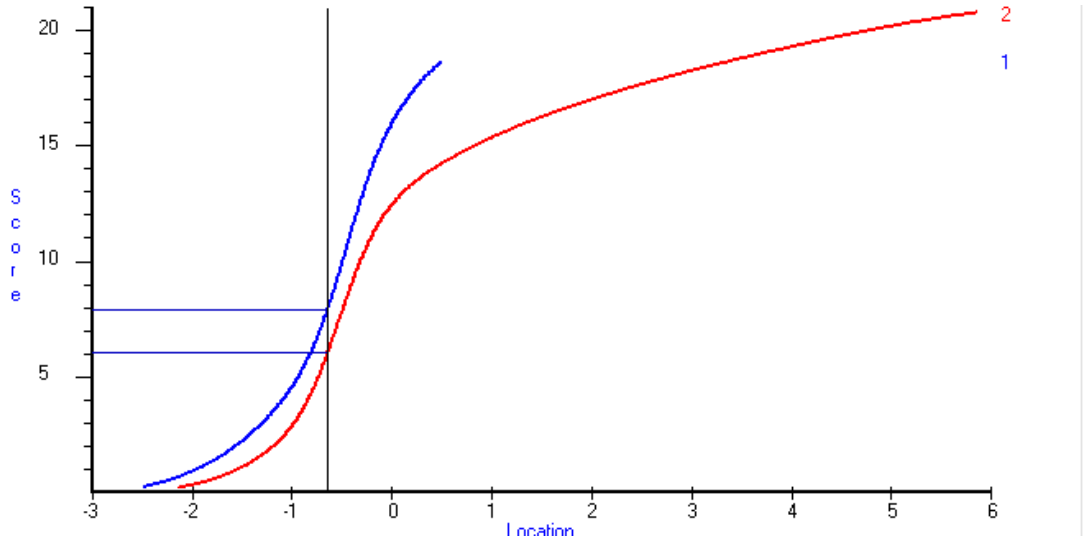

Figure 1. Equating of scores for the HADS-Anxiety (line 1) and HADS-Depression (line 2). Vertical or y axis = score on the HADS-Anxiety or HADS-Depression and horizontal or x axis = location on the logit scale or level of ability. Read a score of 8 across to the HADS-Anxiety (line 1); then read down to HADS-Depression (line 2), and then read back to the raw score, which is 6.

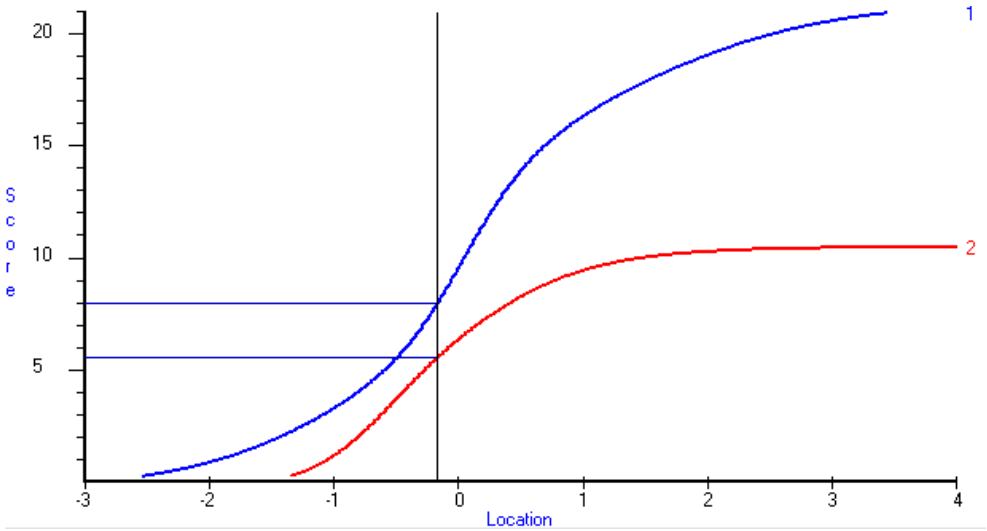

Figure 2. Equating of scores for the HADS-Anxiety (line 1) and DT (line 2). Vertical or y axis = score on the HADS-Anxiety or DT and horizontal or x axis = location on the logit scale or level of ability. Read a score of 8 across to the HADS-Anxiety (line 1); then read down to DT (line 2), and then read back to the raw score, which is 6 (rounded).

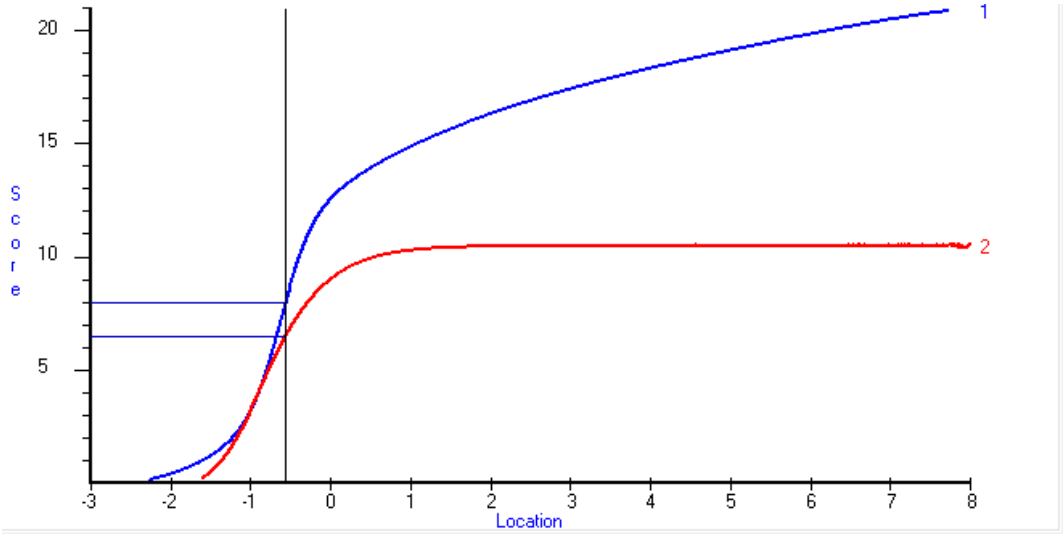

Figure 3. Equating of scores for the HADS-Depression (line 1) and DT (line 2). Vertical or y axis = score on the HADS-Depression or DT and horizontal or x axis = location on the logit scale or level of ability. Read a score of 8 across to the HADS-Depression (line 1); then read down to DT (line 2), and then read back to the raw score, which is 6.5 or 7 rounded.

University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Lambert, SD; Pallant, JF; Clover, K; Britton, B; King, MT; Carter, G

Title:
Using Rasch analysis to examine the distress thermometer's cut-off scores among a mixed group of patients with cancer

Date:
2014-10-01

Citation:
Lambert, S. D., Pallant, J. F., Clover, K., Britton, B., King, M. T. & Carter, G. (2014). Using Rasch analysis to examine the distress thermometer's cut-off scores among a mixed group of patients with cancer. QUALITY OF LIFE RESEARCH, 23 (8), pp.2257-2265. https://doi.org/10.1007/s11136-014-0673-0.

Persistent Link:
http://hdl.handle.net/11343/219784

File Description:
Accepted version