The Validity of the Distress Thermometer in Female Partners of Men with Prostate Cancer

Melissa K. Hyde[1,2], Leah Zajdlewicz[1], Mark Lazenby[3], Jeff Dunn[1,2,4,5], Kirstyn Laurie[1,2], Anthony Lowe[2,7], Suzanne K. Chambers[1,2,4,6,7]

[1] Cancer Council Queensland, Brisbane, Queensland, Australia
[2] Menzies Health Institute Queensland, Griffith University, Gold Coast, Queensland, Australia
[3] School of Nursing, Yale University, New Haven, Connecticut, USA
[4] Institute for Resilient Regions, University of Southern Queensland, Toowoomba, Queensland, Australia
[5] School of Social Science, University of Queensland, St Lucia, Queensland, Australia
[6] Prostate Cancer Foundation of Australia, Sydney, New South Wales, Australia
[7] Exercise Medicine Research Institute, Edith Cowan University, Perth, Western Australia, Australia

* Corresponding author: Professor Suzanne Chambers, Menzies Health Institute Queensland, Griffith University, Gold Coast Campus, Southport, QLD, 4222, Australia. Email: suzanne.chambers@griffith.edu.au

Acknowledgments: Studies were supported by Cancer Council Queensland and the National Health and Medical Research Council (ID496001). We acknowledge Danny Youlden for his statistical support and advice.

Abstract

Female partners of prostate cancer (PCa) survivors experience heightened psychological distress that may be greater than that expressed by PCa patients. However, optimal approaches to detect distressed, or at risk of distress, partners are unclear. This study applied receiver operating characteristics analysis to evaluate diagnostic accuracy, sensitivity and specificity of the Distress Thermometer (DT) compared to widely-used measures of general (Hospital Anxiety and Depression Scale) and cancer-specific (Impact of Events Scale–Revised) distress. Participants were partners of men with localised PCa (recruited around diagnosis) about to undergo or had received surgical treatment (N=189), and partners of men diagnosed with PCa who were 2-4 years post-treatment (N=460). In both studies, diagnostic utility of the DT overall was not optimal. Although area under the curve scores were acceptable (ranges: 0.71 to 0.92 and 0.83 to 0.94 for general and cancer-specific distress, respectively); sensitivity, specificity and optimal DT cut-offs for partner distress varied for general (range: ≥2 to ≥5) and cancer-specific (range: ≥3 to ≥5) distress both across time and between cohorts. Thus, it is difficult to draw firm conclusions about the diagnostic capabilities of the DT for partners or recommend its use in this population. More comprehensive screening measures may be needed to detect partners needing psychological intervention.

Key Words: distress screening; prostate cancer; partners; psychological distress; psychosocial care

Introduction

Globally, prostate cancer (PCa) is the second most common cancer diagnosed in males and the fifth leading cause of cancer death (Torre et al., 2015). Five year prevalence estimates suggest that there are over 3.8 million short- to medium-term PCa survivors globally, with this number expected to increase rapidly in future (Torre et al., 2015). The physical and psychological effects of PCa diagnosis and treatment on these men are well described (Chambers et al., 2017; Chambers, Zajdlewicz, Youlden, Holland, & Dunn, 2014; Hinz et al., 2009; Potosky et al., 2004; Punnen, Cowan, Chan, Carroll, & Cooperberg, 2015), **and research is now focusing on the psychological morbidity experienced by their female partners (Chambers et al., 2017). Specifically, between 36-49% partners of men with prostate cancer report mild to severe anxiety and 9-10% indicate mild to severe depression (Cliff & MacDonagh, 2000) (Chambers et al., 2012), and 20% experience heightened cancer-specific distress (Eton, Lepore, & Helgeson, 2005). Evidence also shows that some partners experience more distress than patients (Cliff & MacDonagh, 2000; Couper et al., 2006; Eton et al., 2005) .** Partner distress can be long-term, with greater distress linked to disease-related (e.g., increased caregiver burden, disease-specific quality of life), individual (e.g., coping style, stress appraisal, physical health), and relationship (e.g., decreased relationship satisfaction, intimacy or functioning) factors (Chambers et al., 2018; Couper et al., 2006; Harden et al., 2013; Hyde et al.; Wootten, Abbott, Farrell, Austin, & Klein, 2014).

In order to facilitate and better target psychological interventions and supportive care in the oncology setting to those who are most likely to benefit, it is necessary to first identify who is distressed **and from this refer to an appropriate level of psychosocial care (Hutchison et al., 2006). This approach ensures valuable and often scarce resources are directed to where they are most needed.** Standards for optimal oncology care advise that

screening for distress and referral to evidence-based services is needed as part of routine care (Holland, Watson, & Dunn, 2011). Ultra-short screening measures (< 5 items, < 2 minutes to complete) are more easily integrated with the demands of clinical care settings (time- and cost-efficient), help to promote universal access to psychological services, and are recommended as an initial step to identify patients who may be distressed (Vodermaier, Linden, & Siu, 2009). The distress thermometer (DT) is widely used with cancer patients as a brief screen to detect overall emotional distress (Lazenby, Tan, Pasacreta, Ercolano, & McCorkle, 2015; Vodermaier et al., 2009). Meta-analysis of validation studies and systematic reviews support an optimal DT cut-off ≥4 to indicate at least moderate distress requiring more in-depth assessment, intervention or referral (Butow et al., 2015; Ma et al., 2014; Mitchell, 2007; Vodermaier et al., 2009). Although use of the DT and its validity is well established for cancer patients including men diagnosed with PCa, (Chambers et al., 2014; Roth et al., 1998), the optimal approach for detecting heightened distress in thee partners of these men is unclear.

A rapid review conducted for the purposes of this study found that quantitative studies with partners or caregivers of adult cancer patients which included the DT as a measure to assess distress (predictor or outcome) typically used a DT cut-off ≥4 (Chambers et al., 2014; Chambers et al., 2012; Chatterton et al., 2015; Fujinami et al., 2015; Howell et al., 2011; Hughes, Sargeant, & Hawkes, 2011; Hutchison et al., 2011; Morris et al., 2015; Trad et al., 2015) or ≥5 (Badr, Gupta, Sikora, & Posner, 2014; Feiten et al., 2013; Goebel, von Harscher, & Mehdorn, 2011; Goldzweig, Rottenberg, Peretz, & Baider, 2015; Halkett et al., 2016; Juarez, Ferrell, Uman, Podnos, & Wagman, 2008; Long et al., 2016; Sklenarova, Haun, et al., 2015; Ugalde, Krishnasamy, & Schofield, 2014; Weide et al., 2012; Zwahlen, Hagenbuch, Jenewein, Carley, & Buchi, 2011). However, systematic reviews and meta-analyses of studies

validating the DT or brief screening measures in the cancer context are silent with regards to the optimal DT cut-off to identify distress in partners/caregivers.

To our knowledge three studies (Bevans et al., 2011; Hawkes, Hughes, Hutchison, & Chambers, 2010; Zwahlen, Hagenbuch, Carley, Recklitis, & Buchi, 2008) have validated the DT in family members or caregivers using receiver operating characteristics (ROC) analysis. Zwahlen et al. (2008) compared the DT to the HADS anxiety and depression subscales and recommended it for use in clinical assessment of family members of outpatients with cancer. A DT cut-off of 3-4 or 4-5 was proposed as optimal for sensitivity (anxiety and depression ≥0.86) and a cut-off of 4-5 for acceptable specificity (anxiety 0.71, depression 0.68). By comparison, Hawkes et al. (2010) compared the DT to the HADS in carers of cancer patients who contacted a cancer helpline and recommended a DT cut-off of ≥6. However, although discriminant ability of the DT (AUC 0.70) and sensitivity (0.77) were acceptable, specificity was poor (0.48). Bevans et al. (2011) compared the DT to the Brief Symptom Inventory-18 in caregivers of patients undergoing allogeneic hematopoietic stem cell transplant. Discriminant ability of the DT was poor (AUC 0.61), and no clear consensus was obtained on an appropriate DT cut-off.

Cancer-specific distress is also an important outcome for caregivers of cancer patients. For example, in a study of 427 female partners of prostate cancer survivors (2-4 years post treatment) who were followed up over 24 months, higher caregiver burden and more threat appraisals were associated with increased distress, anxiety, depression, and cancer-specific distress over time (as measured by the HADS and Impact of Events Scale-Revised [IES-R]) (Chambers et al., 2018). However, there does not appear to be any published studies examining validity of the DT against cancer-specific distress measures such as the IES-R. In sum, there is limited data to guide use of the DT in partners

or carers of cancer patients and no data at all on how useful this scale might be for the partners of men with PCa.

There is a need for guidance regarding diagnostic utility of the DT and its appropriateness for use in oncology and community and primary care settings with partners of cancer patients, and in particular partners of men with PCa. Accordingly, we undertook ROC analyses in two cohorts of women whose partners had been previously diagnosed with PCa. Specifically, we assessed sensitivity and specificity of the DT when compared to general (HADS) and cancer-specific (IES-R) distress, and performance of the DT over time compared to these measures.

Method

Participants

Study 1

Participants were 189 partners of men with localised PCa (recruited around diagnosis) who had chosen or received surgical treatment (Chambers et al., 2008; Chambers et al., 2013). In brief, mean age of partners was 59.8 years (SD = 7.4; range 40.0 to 75.0). Most (96%) partners were married (length of relationship M = 32.5 years, SD = 11.8, range = 0.5 to 51.0) and 30% had children living at home. Approximately half (48%) had a university degree or technical certificate and 40% had a gross household income below AUD$60,000 per annum.

Study 2

Participants were 460 female partners of men previously diagnosed with PCa (2 to 4 years post-treatment), identified through an existing study (Chambers et al., 2017). In brief, partners were mean age 62.8 years (SD = 8.0; range = 29.5-83.0); 96% were married; mean length of relationship was 36.3 years (SD = 12.8; range = 2.1 to 59.5); 21% had children

living at home. Under half (44%) had a university degree or technical certificate; 60% had a gross annual household income below AUD$60,000.

Procedure

Participants gave written informed consent. Griffith University and participating hospitals granted ethics approval. Partners completed the DT (National Comprehensive Cancer Network., Accessed February), the HADS (Zigmond AS & Snaith RP., 1983) at baseline, 3, 6, and 12 months post-diagnosis (Study 1) and the HADS and IES-R (Chambers et al., 2014; Weiss, Marmar, Wilson, & Keane, 1997) at baseline, 6, 12, 18 and 24 months (Study 2).

Measures

Distress thermometer

The DT is widely used as an ultra-brief screening measure to assess general psychological distress in patients diagnosed with cancer (Mitchell, 2007; National Comprehensive Cancer Network., 2013). Partners reported the level of distress they were currently experiencing (within the past week including the day on which they were assessed) on a single 11-point scale, scored 0 *no distress* to 10 *extreme distress*.

Hospital Anxiety and Depression Scale

The 14-item HADS (HADS-T) measured generalized anxiety (HADS-A subscale, 7 items) and depression (HADS-D subscale, 7 items) (Zigmond & Snaith, 1983). Partners rated their experience of anxiety or depression symptoms in the past week on 4-point scales scored 0 (e.g., *not at all*) to 3 (e.g., *very often*); higher scores indicate increased distress (range 0-21 for subscales and 0-42 for total scale). Internal consistency for HADS-A (Study 1: $\alpha = 0.85$ to 0.87; Study 2: $\alpha = 0.87$ to 0.91), HADS-D (Study 1: $\alpha = 0.80$ to 0.84; Study 2: $\alpha = 0.82$ to 0.86), and HADS-T (Study 1: $\alpha = 0.89$ to 0.90; Study 2: $\alpha = 0.90$ to 0.93) was high across all time points in both cohorts. In the current study, cut-off scores ≥8 on HADS-A and HADS-D

subscales (Zigmond AS & Snaith RP., 1983) and ≥11 on HADS-T (Vodermaier & Millman, 2011) **were used to indicate participants with some level of emotional distress.**

Impact of Events Scale-Revised

The 22-item IES-R (Horowitz, Wilner, & Alvarez, 1979; Weiss et al., 1997) measured cancer-specific distress, specifically partner's experience of intrusive or avoidant thinking and hyperarousal regarding the man's PCa in the past week (Study 2 only). Items were scored on 5-point scales from 0 *not at all* to 4 *extremely*; higher scores indicate more cancer-specific distress (range 0 to 88). Internal consistency over time was excellent (α = 0.94 to 0.96). A cut-off score ≥33 was used in the current study to indicate high cancer-specific distress (Creamer, Bell, & Failla, 2003).

Statistical analyses

Receiver operating characteristics (ROC) analysis evaluated diagnostic accuracy of the DT to detect clinical cases across time against the HADS-A, HADS-D, HADS-T and IES-R. ROC analysis generates a curve and the area under the curve (AUC), in this context, represents the diagnostic utility of the DT as compared to the HADS/IES-R across the full range of DT scores (0-10). An AUC of 1 represents perfect agreement between the DT and the HADS/IES-R and an AUC of 0.5 indicates that the probability of the DT detecting a person experiencing distress (as determined by the HADS/IES-R scales) is no greater than chance (Metz, 1978). For the current study, an AUC of 0.60-0.69 is considered to have poor discriminating ability, 0.70-0.79 is fair, 0.80-0.89 is good and 0.90 or above is excellent (Ma et al., 2014; Metz, 1978). Sensitivity (proportion of partners correctly identified as distressed) and specificity values (proportion of partners correctly identified as not distressed) were calculated for each of the DT scale points. The optimal DT thresholds were then selected according to two methods: the Youden Index (the cut-off is selected as the point on the AUC that is farthest from chance, corresponding to the maximum value of (sensitivity + specificity

− 1) (Youden., 1950), and the point on the ROC closest to (0,1) indicating perfect classification of partners with and without clinical distress. Table 1 shows the optimal DT cut-offs according to the two selection methods along with the AUC, sensitivity, specificity, the percentage correctly classified, and the positive and negative likelihood ratios.

Results

Study 1

Psychological distress

Distress thermometer scores at baseline (*M* = 2.79, *SD* = 2.66), 3 months (*M* = 1.87, *SD* = 2.47), 6 months (*M* = 1.72, *SD* = 2.14) and 12 months (*M* = 1.52, *SD* = 2.06) showed that on average, partner distress decreased after the first 6 months. The proportion of partners who would be considered distressed using DT cut-offs suggested for caregivers of patients with cancer in prior research was identified. Based on a DT cut-off ≥4 (Zwahlen et al., 2008), 23.3% of partners were distressed at baseline, and this reduced to 14.5% and 14.2% at 3 and 6 months (by which time most men had received treatment), respectively, and decreased further to 11.5% at 12 months. With a DT cut-off ≥5 (Zwahlen et al., 2008), 18.5% were identified as distressed at baseline and the proportion of partners who were distressed decreased over time (3 months 12.6%; 6 months 9.0%; 12 months 7.4%). This pattern was also consistent with a DT cut-off ≥6 (Hawkes et al., 2010), with 12.7% of partners identified as distressed at baseline with subsequent declines in these numbers (3 months 8.2%; 6 months 5.2%; 12 months 4.1%).

Diagnostic accuracy of the DT compared to the HADS

Comparing the DT to HADS-T at baseline, the AUC was 0.84 (95% CI 0.78-0.90) with an optimal DT cut-off of ≥4 according to both classification methods (Table 1). The corresponding values for ROC curves for the HADS-T at 3, 6 and 12 months remained fairly

stable, with a decrease in the threshold to ≥3 at 3 and 6 months and then to ≥2 at 12 months. Sensitivity of the DT compared with HADS-T across the time points was low (61% to 69%), however, specificity was higher (73% to 88%). A similar pattern of results to the HADS-T were noted for the HADS-A from baseline to 12 months, except for 2 different optimal DT cut-off points being identified at 3 months of ≥3 (Youden Index) and ≥2 (point on the ROC closest to (0,1)). Both cut-off points had low sensitivity (61.4% and 68.2%) and higher specificity (83.5% and 73.0%).

A different pattern was evident for the HADS-D. At baseline, the AUC was 0.84 (95% CI 0.71-0.96) with a DT cut-off of ≥5 maximising sensitivity (87.5%) and specificity (82.7%). At 3 months, the accuracy of the DT relative to the HADS-D decreased to 0.71 (95% CI 0.54-0.88) and optimal cut-offs were determined as ≥5 (Youden Index) or ≥3 (the point on the ROC closest to (0,1)). Both cut-offs had very low sensitivity (50.0% and 58.3%), however, specificity was higher (88.4% and 73.5%). Diagnostic accuracy of the DT compared to the HADS-D at 3 months was similar to the 6 month and 12 month time points, although the optimal DT cut-offs lowered to ≥2 and ≥3, respectively.

Study 2

Psychological distress

Mean scores on the DT suggest that partner distress remained relatively unchanged from baseline (*M* = 1.91, *SD* = 2.40), 6 months (*M* = 1.94, *SD* = 2.44), 12 months (*M* = 1.95, *SD* = 2.51), and 18 months (*M* = 1.99, *SD* = 2.47), with a slight decline at 24 months (*M* = 1.77, *SD* = 2.42). Using the DT cut-offs suggested for caregivers in prior research, a cut-off ≥4 (Zwahlen et al., 2008) identified18.0% of partners as distressed at baseline and this proportion remained relatively stable over time (6 months 17.5%, 12 months 17.2%, 18 months 18.0%, 24 months 16.7%). A DT cut-off ≥5 (Zwahlen et al., 2008) classified

approximately 10% of partners as distressed over time (baseline 9.8%, 6 months 11.7%, 12 months 11.1%, 18 months 10.8%, 24 months 10.6%); and a DT cut-off ≥6 (Hawkes et al., 2010) identified less than 10% of partners as distressed (baseline 6.1%, 6 months 7.3%, 12 months 8.6%, 18 months 8.0%, 24 months 7.0%).

Diagnostic accuracy of the DT compared to the HADS

For the HADS-T at baseline, a cut-off score of ≥2 yielded the optimal ratio of sensitivity (82.7%) and specificity (75.3%) with an AUC of 0.85 (95% CI 0.81-0.89). Accuracy of the DT compared to the HADS-T remained high (AUCs ≥0.85) at 6, 12 and 18 months, and decreased slightly at 24 months (AUC = 0.81, 95% CI 0.76-0.87). However, sensitivity reduced at 6, 12, and 18 months and specificity increased in the same time frame (Table 1). At the 24-month assessment, sensitivity (77.1%) and specificity (73.0%) were more balanced. The optimal DT cut-off was ≥3 at 6, 12, and 18 months and reduced to ≥2 at 24 months.

Compared to the HADS-A subscale, the DT initially showed a cut-point ≥3 at baseline (and again at 12 months) as providing the best balance of sensitivity (73.1%) and specificity (79.8%). However, the optimal DT cut-point reduced to ≥2 at 6, 18, and 24 months (Table 1). An alternative DT cut-off ≥1 was suggested at 24 months (Youden) and although sensitivity was high (90.2%), specificity was poor (59.2%). All AUCs were >0.80 supporting good accuracy overall across time for the DT compared to the anxiety subscale (Table 1).

The AUC score at baseline (0.87, 95% CI 0.82-0.92) showed good accuracy for the DT as compared to the HADS-D subscale, with a cut-off ≥4 providing optimal balance between sensitivity (79.5%) and specificity (82.3%). At 6, 12, and 24 months, the DT cut-off

providing the best balance for sensitivity and specificity reduced to ≥3 (ROC (0,1)) (Table 1). A cut-point ≥3 (Youden) was also identified for 18 months (AUC = 0.92, 95% CI 0.88-0.95), however a cut-off ≥4 (ROC (0,1)) gave a better balance between sensitivity (86.2%) and specificity (82.8%). For all remaining time points, the DT showed good accuracy with AUCs ranging from 0.81 (6 months) to 0.89 (24 months) (Table 1).

Diagnostic accuracy of the DT compared to the IES-R

At baseline, the AUC was 0.83 (95% CI 0.75-0.91) with an optimal DT cut-off of ≥3 (Table 1). Although accuracy of the DT across time remained high (AUCs ranged from 0.87-0.91), the optimal DT cut-off was not consistent over time. At 6 months a DT cut-off ≥5 provided a good balance between sensitivity (85.0%) and specificity (85.8%), and at 12 months this reduced to ≥3 (Youden Index) or ≥4 (ROC (0,1)); increased to ≥5 (Youden Index) or ≥4 (ROC (0,1)) at 18 months; and remained at ≥5 at 24 months. Over time, depending on the optimal DT cut-off identified, sensitivity and specificity values ranged from 71% to 95% and 71% to 87%, respectively.

Discussion

In the current study, overall diagnostic accuracy of the DT in female partners of men with PCa was inconsistent. Although AUC scores suggested that performance of the DT compared to the HADS (AUC range 0.71 to 0.92) and IES-R (AUC range 0.83 to 0.94) was acceptable, other diagnostic indicators varied and were context-dependent. Specifically, when performance of the DT was compared to the HADS within and between studies, optimal DT cut-offs to detect partner distress ranged from ≥2 to ≥5; against the IES-R, optimal DT cut-offs ranged from ≥3 to ≥5; with no clear pattern emerging to indicate a single optimal DT cut-point. Close to the time of the man's diagnosis and treatment, the sensitivity of the DT in detecting partners' distress was consistently lower than specificity, regardless of the DT cut-

off used. When men were beyond treatment and well into survivorship, sensitivity was either higher than specificity or each value was more equally balanced. In the context of screening for a common condition such as distress, high sensitivity is needed to avoid overlooking partners who are in fact distressed (Youngstrom, 2014). However, unless the measure has reasonable specificity, there is a risk of over-diagnosis of distress resulting in unnecessary treatment or referral and impact on already limited availability of supportive care professionals and resources in clinical care. Thus, use of the DT alone to detect female partners in need of further assessment or psychological intervention appears problematic and more comprehensive screening measures may be needed.

Ultra-short screening measures with 2 to 4 items tend to have better sensitivity than single item measures (Mitchell & Coyne, 2007) although, similar to the DT, specificity may be moderate and evaluation with a more in-depth measure required (Lazenby et al., 2015; Mitchell & Coyne, 2007). The four item Patient Health Questionnaire (PHQ-4) is an ultra-short measure assessing anxiety (Generalised Anxiety Disorder [GAD]-2) and depression (PHQ-2) with acceptable psychometric properties that appears promising for use in primary care settings with patients (Kroenke, Spitzer, Williams, & Löwe, 2009) and warrants further investigation in cancer settings (Pirl et al., 2014). The PHQ-4 is also increasingly being utilised in cross-sectional or longitudinal predictive studies to measure distress as a predictor or outcome in partners (e.g., Ernst et al., 2017; Haun et al., 2014; Pankrath et al., 2016) and caregivers (e.g., Oh, 2017; Sklenarova, Krümpelmann, et al., 2015) of cancer patients. However, no studies have examined the diagnostic utility of the PHQ-4 compared to other ultra-short or more in-depth screening measures such as the HADS in partners or caregivers of cancer patients. This is an area worthy of further investigation.

In some studies of patients with cancer, distress as measured by the DT has aligned more closely with the anxiety subscale of the HADS (Gil et al., 2005). To address this,

researchers have proposed the use of additional emotion thermometers (e.g., mood thermometer) to supplement the DT and increase sensitivity to better detect multiple domains of distress (e.g., depression) (Gil et al., 2005; Mitchell, 2007; Vodermaier et al., 2009). For instance, Mitchell and colleagues (Mitchell, Baker-Glenn, Granger, & Symonds, 2010) proposed the use of a 5-item emotion thermometers tool covering distress, anxiety, depression, anger and need for help and showed that sensitivity of the DT in detecting distressed patients with cancer can be improved by including these additional thermometers (Mitchell, Baker-Glenn, Park, Granger, & Symonds, 2010). For partners or caregivers, there may also be other dimensions of distress such as caregiver burden that warrant inclusion. However, as is clear from the current study, suitable approaches for patients may not transfer directly to partners and a new approach may be needed potentially building on the work of Mitchell and colleagues (Mitchell, Baker-Glenn, Granger, et al., 2010; Mitchell, Baker-Glenn, Park, et al., 2010) to incorporate partner- or caregiver-specific issues.

Overall, our results are not consistent with other validations of the DT with caregivers/family members in cancer (Hawkes et al., 2010; Zwahlen et al., 2008). Of note, these previous studies relied on single sample, cross-sectional designs within which temporal and contextual variations will not be elucidated. A strength of the current study is the inclusion of two cohorts of partners prospectively assessed at different time points in the cancer trajectory. A limitation of the current study is the small number of female partners who were identified as depressed (~8%) resulting in wide confidence intervals around the AUC scores and potentially increasing uncertainty regarding optimal thresholds (Chambers et al., 2014). As noted, there may also be other dimensions of distress that are more salient for partners (e.g., caregiver burden, Adelman, Tmanova, Delgado, Dion, & Lachs, 2014) but were not assessed in the current study. Future research could examine alternative or a broader

range of criterion measures, including those developed specifically for partners/caregivers, against which to compare suitability or otherwise of the DT.

In conclusion, when compared to well-used and validated measures of general and cancer-specific distress overall diagnostic performance of the DT in female partners of men with PCa was not strong. Optimal cut-offs and sensitivity and specificity of the DT varied across time and between cohorts with no clear pattern emerging. Thus, at this point in time, it is not possible to draw firm conclusions about optimal DT cut-offs or recommend use of the DT to detect distress in partners of men diagnosed with PCa. Further work is required to identify the most suitable screening measure for female partners/carers in need of psychological intervention **that may then be implemented in acute, community and primary care settings.**

References

Adelman, R. D., Tmanova, L. L., Delgado, D., Dion, S., & Lachs, M. S. (2014). Caregiver burden: a clinical review. *Jama, 311*(10), 1052-1060.

Badr, H., Gupta, V., Sikora, A., & Posner, M. (2014). Psychological distress in patients and caregivers over the course of radiotherapy for head and neck cancer. *Oral oncology, 50*(10), 1005-1011.

Bevans, M., Wehrlen, L., Prachenko, O., Soeken, K., Zabora, J., & Wallen, G. R. (2011). Distress screening in allogeneic hematopoietic stem cell (HSCT) caregivers and patients. *Psycho‐Oncology, 20*(6), 615-622.

Butow, P., Price, M. A., Shaw, J. M., Turner, J., Clayton, J. M., Grimison, P., . . . Kirsten, L. (2015). Clinical pathway for the screening, assessment and management of anxiety and depression in adult cancer patients: Australian guidelines. *Psycho‐Oncology, 24*(9), 987-1001.

Chambers SK, Hyde MK, Smith DP, Hughes S, Yuill S, Egger S, . . . Dunn J. (2017). A systematic review of psychological interventions for prostate cancer survivors and their partners: Clinical and research implications. *Psycho-Oncology*. doi:10.1002/pon.4431

Chambers SK, Ng SK, Baade P, Aitken JF, Hyde MK, Wittert G, . . . Dunn J. (2017). Trajectories of quality of life, life satisfaction, and psychological adjustment after prostate cancer. *Psycho‐Oncology*. doi:10.1002/pon.4342

Chambers SK, Zajdlewicz L, Youlden DR, Holland JC, & Dunn J. (2014). The validity of the distress thermometer in prostate cancer populations. *Psycho‐Oncology, 23*(2), 195-203. doi:10.1002/pon.3391

Chambers, S. K., Girgis, A., Occhipinti, S., Hutchison, S., Turner, J., McDowell, M., . . . Dunn, J. (2014). *A randomized trial comparing two low-intensity psychological interventions for distressed patients with cancer and their caregivers.* Paper presented at the Oncology nursing forum.

Chambers, S. K., Girgis, A., Occhipinti, S., Hutchison, S., Turner, J., Morris, B., & Dunn, J. (2012). Psychological distress and unmet supportive care needs in cancer patients and carers who contact cancer helplines. *European Journal of Cancer Care, 21*(2), 213-223.

Chambers, S. K., Schover, L., Halford, K., Clutton, S., Ferguson, M., Gordon, L., . . . Dunn, J. (2008). ProsCan for Couples: randomised controlled trial of a couples-based sexuality intervention for men with localised prostate cancer who receive radical prostatectomy. *BMC cancer, 8*(1), 226.

Chambers, S. K., Schover, L., Nielsen, L., Halford, K., Clutton, S., Gardiner, R. A., . . . Occhipinti, S. (2013). Couple distress after localised prostate cancer. *Supportive Care in Cancer, 21*(11), 2967-2976.

Chatterton, M. L., Chambers, S., Occhipinti, S., Girgis, A., Dunn, J., Carter, R., . . . Mihalopoulos, C. (2015). Economic evaluation of a psychological intervention for high distress cancer patients and carers: costs and quality‐ adjusted life years. *Psycho‐Oncology*.

Cliff AM, MacDonagh RP. Psychosocial morbidity in prostate cancer: II. A comparison of patients and partners. BJU Int. 2000 Nov;86(7):834-9. PMID: 11069410. doi: bju914 [pii].

Couper J, Bloch S, Love A, Macvean M, Duchesne GM, & Kissane D. (2006). Psychosocial adjustment of female partners of men with prostate cancer: a review of the literature. *Psycho‐Oncology, 15*(11), 937-953. doi:10.1002/pon.1031

Creamer, M., Bell, R., & Failla, S. (2003). Psychometric properties of the impact of event scale—revised. *Behaviour research and therapy, 41*(12), 1489-1496.

Ernst, J., Hinz, A., Niederwieser, D., Döhner, H., Hönig, K., Vogelhuber, M., . . . Weissflog, G. (2017). Dyadic coping of patients with hematologic malignancies and their partners and its relation to quality of life–a longitudinal study. *Leukemia & Lymphoma, 58*(3), 655-665.

Eton DT, Lepore SJ, Helgeson VS. Psychological distress in spouses of men treated for early-stage prostate carcinoma. Cancer. 2005 Jun 1;103(11):2412-8. PMID: 15858824. doi: 10.1002/cncr.21092.

Feiten, S., Friesenhahn, V., Heymanns, J., Kleboth, K., Köppler, H., Mergenthaler, U., . . . Weide, R. (2013). Psychosocial distress in caregivers of patients with a metastatic solid tumor in routine care: a survey in a community based oncology group practice in Germany. *Cancer and Clinical Oncology, 2*(2), 1.

Fujinami, R., Sun, V., Zachariah, F., Uman, G., Grant, M., & Ferrell, B. (2015). Family caregivers' distress levels related to quality of life, burden, and preparedness. *Psycho-Oncology, 24*(1), 54-62.

Gil, F., Grassi, L., Travado, L., Tomamichel, M., Gonzalez, J. R., & Group, S. (2005). Use of distress and depression thermometers to measure psychosocial morbidity among southern European cancer patients. *Supportive Care in Cancer, 13*(8), 600-606.

Goebel, S., von Harscher, M., & Mehdorn, H. M. (2011). Comorbid mental disorders and psychosocial distress in patients with brain tumours and their spouses in the early treatment phase. *Supportive Care in Cancer, 19*(11), 1797-1805.

Goldzweig, G., Rottenberg, Y., Peretz, T., & Baider, L. (2015). Silent Partners to Cancer Patients: Formal Caregivers and Oncologists. *Journal of Cancer Education, 30*(4), 704-710.

Halkett, G. K., Lobb, E. A., Shaw, T., Sinclair, M. M., Miller, L., Hovey, E., & Nowak, A. K. (2016). Distress and psychological morbidity do not reduce over time in carers of patients with high-grade glioma. *Supportive Care in Cancer*, 1-7.

Harden, J. K., Sanda, M. G., Wei, J. T., Yarandi, H., Hembroff, L., Hardy, J., . . . Group, P. C. S. (2013). Partners' long-term appraisal of their caregiving experience, marital satisfaction, sexual satisfaction, and quality of life 2 years after prostate cancer treatment. *Cancer nursing, 36*(2).

Haun, M. W., Sklenarova, H., Villalobos, M., Thomas, M., Brechtel, A., Löwe, B., . . . Hartmann, M. (2014). Depression, anxiety and disease-related distress in couples affected by advanced lung cancer. *Lung Cancer, 86*(2), 274-280.

Hawkes, A. L., Hughes, K. L., Hutchison, S. D., & Chambers, S. K. (2010). Feasibility of brief psychological distress screening by a community-based telephone helpline for cancer patients and carers. *BMC cancer, 10*(1), 14.

Hinz, A., Krauss, O., Stolzenburg, J.-U., Schwalenberg, T., Michalski, D., & Schwarz, R. (2009). *Anxiety and depression in patients with prostate cancer and other urogenital cancer: a longitudinal study.* Paper presented at the Urologic Oncology: Seminars and Original Investigations.

Holland J, Watson M, & Dunn J. (2011). The IPOS new international standard of quality cancer care: Integrating the psychosocial domain into routine care. *Psycho-Oncology*, *20*(7), 677-680. doi:10.1002/pon.1978

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: a measure of subjective stress. *Psychosomatic medicine, 41*(3), 209-218.

Howell, D., Marshall, D., Brazil, K., Taniguchi, A., Howard, M., Foster, G., & Thabane, L. (2011). A shared care model pilot for palliative home care in a rural area: impact on symptoms, distress, and place of death. *Journal of pain and symptom management, 42*(1), 60-75.

Hughes, K. L., Sargeant, H., & Hawkes, A. L. (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), 46.

Hutchison, S. D., Sargeant, H., Morris, B. A., Hawkes, A. L., Clutton, S., & Chambers, S. K. (2011). A community- based approach to cancer counselling for patients and carers: a preliminary study. *Psycho- Oncology, 20*(8), 897-901.

Hutchison SD, Steginga SK, Dunn J. The tiered model of psychosocial intervention in cancer: a community based approach. Psychooncology. 2006 Jun;15(6):541-6. PMID: 16331595. doi: 10.1002/pon.973.

Hyde MK, Legg M, Lepore S, Ugalde A, Gough K, Zajdlewicz L, . . . Chambers SK. (2018). Trajectories of distress in female partners of men diagnosed with prostate cancer.

Juarez, G., Ferrell, B., Uman, G., Podnos, Y., & Wagman, L. D. (2008). Distress and quality of life concerns of family caregivers of patients undergoing palliative surgery. *Cancer nursing, 31*(1), 2-10.

Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: the PHQ–4. *Psychosomatics, 50*(6), 613-621.

Lazenby, M., Tan, H., Pasacreta, N., Ercolano, E., & McCorkle, R. (2015). The five steps of comprehensive psychosocial distress screening. *Current oncology reports, 17*(5), 1-5.

Long, A., Halkett, G. K., Lobb, E. A., Shaw, T., Hovey, E., & Nowak, A. K. (2016). Carers of patients with high-grade glioma report high levels of distress, unmet needs, and psychological morbidity during patient chemoradiotherapy. *Neuro-oncology practice, 3*(2), 105-112.

Ma, X., Zhang, J., Zhong, W., Shu, C., Wang, F., Wen, J., . . . Liu, L. (2014). The diagnostic role of a short screening tool—the distress thermometer: a meta-analysis. *Supportive Care in Cancer, 22*(7), 1741-1755.

Metz, C. E. (1978). *Basic principles of ROC analysis.* Paper presented at the Seminars in nuclear medicine.

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, 25*(29), 4670-4681.

Mitchell, A. J., Baker- Glenn, E. A., Granger, L., & Symonds, P. (2010). Can the Distress Thermometer be improved by additional mood domains? Part I. Initial validation of the Emotion Thermometers tool. *Psycho- Oncology, 19*(2), 125-133.

Mitchell, A. J., Baker- Glenn, E. A., Park, B., Granger, L., & Symonds, P. (2010). Can the Distress Thermometer be improved by additional mood domains? Part II. What is the optimal combination of Emotion Thermometers? *Psycho- Oncology, 19*(2), 134-140.

Mitchell, A. J., & Coyne, J. C. (2007). Do ultra-short screening instruments accurately detect depression in primary care? *British Journal of General Practice, 57*(535), 144-151.

Morris, B. A., Thorndike, F. P., Ritterband, L. M., Glozier, N., Dunn, J., & Chambers, S. K. (2015). Sleep disturbance in cancer patients and caregivers who contact telephone-based help services. *Supportive Care in Cancer, 23*(4), 1113-1120.

National Comprehensive Cancer Network. (Accessed February 2016). Distress Management Version 2.2013. *http://www.nccn.org/professionals/physician_gls/PDF/distress.pdf*.

Oh, Y. S. (2017). Communications with health professionals and psychological distress in family caregivers to cancer patients: A model based on stress-coping theory. *Applied Nursing Research, 33*, 5-9.

Pankrath, A. L., Weißflog, G., Mehnert, A., Niederwieser, D., Döhner, H., Hönig, K., . . . Ernst, J. (2016). The relation between dyadic coping and relationship satisfaction in couples dealing with haematological cancer. *European Journal of Cancer Care*.

Pirl, W. F., Fann, J. R., Greer, J. A., Braun, I., Deshields, T., Fulcher, C., . . . Lazenby, M. (2014). Recommendations for the implementation of distress screening programs in cancer centers: report from the American Psychosocial Oncology Society (APOS), Association of Oncology Social Work (AOSW), and Oncology Nursing Society (ONS) joint task force. *Cancer, 120*(19), 2946-2954.

Potosky, A. L., Davis, W. W., Hoffman, R. M., Stanford, J. L., Stephenson, R. A., Penson, D. F., & Harlan, L. C. (2004). Five-year outcomes after prostatectomy or radiotherapy for prostate cancer: the prostate cancer outcomes study. *Journal of the National Cancer Institute, 96*(18), 1358-1367.

Punnen, S., Cowan, J. E., Chan, J. M., Carroll, P. R., & Cooperberg, M. R. (2015). Long-term health-related quality of life after primary treatment for localized prostate cancer: results from the CaPSURE registry. *European urology, 68*(4), 600-608.

Roth, A. J., Kornblith, A. B., Batel- Copel, L., Peabody, E., Scher, H. I., & Holland, J. C. (1998). Rapid screening for psychologic distress in men with prostate carcinoma. *Cancer, 82*(10), 1904-1908.

Sklenarova, H., Haun, M., Krümpelmann, A., Friederich, H. C., Huber, J., Thomas, M., . . . Herzog, W. (2015). Psychometric evaluation of the German Version of the Supportive Care Needs Survey for Partners and Caregivers (SCNS- P&C- G) of cancer patients. *European Journal of Cancer Care, 24*(6), 884-897.

Sklenarova, H., Krümpelmann, A., Haun, M. W., Friederich, H. C., Huber, J., Thomas, M., . . . Hartmann, M. (2015). When do we need to care about the caregiver? Supportive care needs, anxiety, and depression among informal caregivers of patients with cancer and cancer survivors. *Cancer, 121*(9), 1513-1519.

Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, & Jemal A. (2015). Global cancer statistics 2012. *CA: A Cancer Journal for Clinicians, 65*(2), 87-108. doi:10.3322/caac.21262

Trad, W., Koh, E.-S., Daher, M., Bailey, A., Kastelan, M., Legge, D., . . . Hovey, E. (2015). Screening for psychological distress in adult primary brain tumor patients and caregivers: considerations for cancer care coordination. *Frontiers in oncology, 5*.

Ugalde, A., Krishnasamy, M., & Schofield, P. (2014). The relationship between self-efficacy and anxiety and general distress in caregivers of people with advanced cancer. *Journal of palliative medicine, 17*(8), 939-941.

Vodermaier, A., Linden, W., & Siu, C. (2009). Screening for emotional distress in cancer patients: a systematic review of assessment instruments. *Journal of the National Cancer Institute, 101*(21), 1464-1488.

Vodermaier, A., & Millman, R. D. (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.

Weide, R., Feiten, S., Friesenhahn, V., Heymanns, J., Kleboth, K., Mergenthaler, U., . . . Köppler, H. (2012). Identifying Caregivers and Their Meaning for Patients with Metastatic Solid Tumours in Routine Care: A Survey at a Community-Based Oncology Group Practice in Germany. *Cancer and Clinical Oncology, 1*(1), 41.

Weiss, D., Marmar, C., Wilson, J., & Keane, T. (1997). The Impact of Events Scale-Revised. In Wilson JP & Keane TM (Eds.), *Assessing psychological trauma and PTSD* (pp. 399-411). New York, NY: Guilford Press.

Wootten AC, Abbott JM, Farrell A, Austin DW, & Klein B. (2014). Psychosocial interventions to support partners of men with prostate cancer: a systematic and critical review of the literature. *Journal of Cancer Survivorship, 8*(3), 472-484. doi:10.1007/s11764-014-0361-7

Youden WJ. (1950). Index for rating diagnostic tests. *Cancer, 3*, 32-35.

Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of pediatric psychology, 39*(2), 204-221.

Zigmond AS, & Snaith RP. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*(6), 361-370.

Zwahlen, D., Hagenbuch, N., Carley, M. I., Recklitis, C. J., & Buchi, S. (2008). Screening cancer patients' families with the distress thermometer (DT): a validation study. *Psycho‐Oncology, 17*(10), 959-966.

Zwahlen, D., Hagenbuch, N., Jenewein, J., Carley, M. I., & Buchi, S. (2011). Adopting a family approach to theory and practice: measuring distress in cancer patient–partner dyads with the distress thermometer. *Psycho‐Oncology, 20*(4), 394-403.

Table 1. Summary of area under the ROC curve analyses comparing the DT to the HADS and IES-R

Comparative scale[a]	Assessment point (months)	Cases/Cohort (n)	Area under ROC curve (95% CI)	Optimum cut-point for DT[b]	Sensitivity (%)	Specificity (%)	Correctly classified (%)	Positive likelihood ratio[d]	Negative likelihood ratio[e]
Study 1 (N = 189)									
HADS-A	0 (baseline)	68/189	0.83 (0.77-0.89)	≥ 4 [b]	66.18	88.43	80.42	5.72	0.38
	0	68/189	0.83 (0.77-0.89)	≥ 3 [c]	77.94	75.21	76.19	3.14	0.29
	3	44/159	0.78 (0.70-0.86)	≥ 3 [b]	61.36	83.48	77.36	3.71	0.46
	3	44/159	0.78 (0.70-0.86)	≥ 2 [c]	68.18	73.04	71.70	2.53	0.44
	6	30/155	0.82 (0.73-0.91)	≥ 3	70.00	82.40	80.00	3.98	0.36
	12	30/148	0.79 (0.70-0.88)	≥ 2	73.33	72.03	72.30	2.62	0.37
HADS-D	0 (baseline)	16/189	0.84 (0.71-0.96)	≥ 5	87.50	82.66	83.07	5.05	0.15
	3	12/159	0.71 (0.54-0.88)	≥ 5 [b]	50.00	88.44	85.53	4.32	0.57
	3	12/159	0.71 (0.54-0.88)	≥ 3 [c]	58.33	73.47	72.33	2.20	0.57
	6	6/155	0.73 (0.51-0.95)	≥ 2	83.33	63.09	63.87	2.26	0.26
	12	11/148	0.76 (0.59-0.93)	≥ 3	63.64	79.56	78.38	3.11	0.46
HADS-T	0 (baseline)	68/189	0.84 (0.78-0.90)	≥ 4	66.18	88.43	80.42	5.72	0.38
	3	44/159	0.78 (0.70-0.86)	≥ 3	61.36	83.48	77.36	3.71	0.46
	6	35/155	0.81 (0.72 -0.90)	≥ 3	68.57	84.17	80.65	4.33	0.37
	12	36/148	0.76 (0.66 -0.86)	≥ 2	69.44	73.21	72.30	2.59	0.42
Study 2 (N = 460)									
HADS-A	0 (baseline)	104/425	0.81 (0.76-0.86)	≥ 3	73.08	79.75	78.12	3.61	0.34
	6	97/393	0.83 (0.79-0.88)	≥ 2	82.47	69.59	72.77	2.71	0.25
	12	88/360	0.85 (0.80-0.90)	≥ 3	76.14	82.35	80.83	4.31	0.29
	18	82/350	0.85 (0.81-0.90)	≥ 2	85.37	71.27	74.57	2.97	0.21
	24	82/359	0.81 (0.76-0.86)	≥ 1 [b]	90.24	59.21	66.30	2.21	0.16
	24	82/359	0.81 (0.76-0.86)	≥ 2 [c]	76.83	72.20	73.26	2.76	0.32
HADS-D	0 (baseline)	39/423	0.87 (0.82-0.92)	≥ 4	79.49	82.29	82.03	4.49	0.25
	6	42/393	0.81 (0.73-0.88)	≥ 3	83.33	75.78	76.59	3.44	0.22

Comparative scale[a]	Assessment point (months)	Cases/Cohort (n)	Area under ROC curve (95% CI)	Optimum cut-point for DT[b]	Sensitivity (%)	Specificity (%)	Correctly classified (%)	Positive likelihood ratio[d]	Negative likelihood ratio[e]
	12	35/361	0.85 (0.79-0.91)	≥ 2 [b]	94.29	64.11	67.04	2.63	0.09
	12	35/361	0.85 (0.79-0.91)	≥ 3 [c]	82.86	73.31	74.24	3.10	0.23
	18	29/349	0.92 (0.88-0.95)	≥ 3 [b]	96.55	73.44	75.36	3.63	0.05
	18	29/349	0.92 (0.88-0.95)	≥ 4 [c]	86.21	82.81	83.09	5.02	0.17
	24	28/357	0.89 (0.82-0.95)	≥ 2 [b]	96.43	66.26	68.63	2.86	0.05
	24	28/357	0.89 (0.82-0.95)	≥ 3 [c]	78.57	77.51	77.59	3.49	0.28
HADS-T	0 (baseline)	110/422	0.85 (0.81-0.89)	≥ 2	82.73	75.32	77.25	3.35	0.23
	6	107/391	0.85 (0.80-0.89)	≥ 3	72.90	85.21	81.84	4.93	0.32
	12	98/360	0.85 (0.80-0.89)	≥ 3	75.51	84.35	81.94	4.83	0.29
	18	93/349	0.88 (0.84-0.92)	≥ 3	78.49	84.38	82.81	5.02	0.25
	24	83/357	0.81 (0.76-0.87)	≥ 2	77.11	72.99	73.95	2.86	0.31
IES-R	0 (baseline)	27/418	0.83 (0.75-0.91)	≥ 3	85.19	70.59	71.53	2.90	0.21
	6	20/387	0.91 (0.87-0.95)	≥ 5	85.00	85.83	85.79	6.00	0.17
	12	22/361	0.91 (0.87-0.95)	≥ 3 [b]	95.45	71.98	73.41	3.41	0.06
	12	22/361	0.91 (0.87-0.95)	≥ 4 [c]	81.82	81.42	81.44	4.40	0.22
	18	21/345	0.87 (0.79-0.95)	≥ 5 [b]	71.43	86.11	85.22	5.14	0.33
	18	21/345	0.87 (0.79-0.95)	≥ 4 [c]	76.19	80.86	80.58	3.98	0.29
	24	17/356	0.94 (0.90-0.98)	≥ 5	88.24	87.02	87.08	6.80	0.14

Note. DT = Distress Thermometer; HADS-A = Hospital Anxiety and Depression Scale-Anxiety subscale; HADS-D = Hospital Anxiety and Depression Scale-Depression subscale; HADS-T = Hospital Anxiety and Depression Scale-Total score; IES-R = Impact of Events Scale-Revised total score.

a. Caseness defined as ≥ 8 on HADS-A or HADS-D subscales and ≥ 11 on the HADS-T; Caseness defined as ≥ 33 on IES-R.
b. The optimal DT cut-off is given by the maximum value of the Youden index which measures the vertical distance from the line of equality to the ROC curve.
c. The optimal DT cut-off is given by the point that is closest to the ROC (0, 1).
d. Positive likelihood ratio=Sensitivity/(1–Specificity).
e. Negative likelihood ratio=(1–Sensitivity)/Specificity