The diagnostic role of a short screening tool – the distress thermometer: a meta-analysis


CRD summary
This review concluded that the distress thermometer had good accuracy in detecting distress in cancer patients; a cut-off score of 4 (where 10 indicated extreme distress) was considered optimal. Although the evidence base was large, the lack of a consistent and reliable reference standard suggests the authors’ conclusions should be considered carefully as the test accuracy may be overestimated.

Authors' objectives
To assess the diagnostic accuracy of the distress thermometer in detecting distress in cancer patients.

Searching
PubMed and EMBASE were searched from 1997 up to September 2013 with no language restrictions. Search terms were reported.

Study selection
Eligible for inclusion in the review were studies that assessed the diagnostic accuracy of the distress thermometer to detect psychological disorders in patients with cancer. Eligible studies had to compare accuracy of the thermometer to a reference standard and present sufficient data to construct a 2x2 table of diagnostic accuracy. Abstracts were excluded from the review.

Included studies were published from 2003 to 2013; they were conducted in 20 different countries (four studies in the UK). The mean age of patients ranged from 23.5 to 70 years. Most studies included multiple types of cancer. Patients were survivors and/or undergoing active treatment, and/or at the end of life (cancer trajectory). Cut-off scores for the distress thermometer varied across studies, ranging from 2 to 7 out of 10 (10 indicated extreme distress). Ten different reference standards were used; versions of the Hospital Anxiety and Depression Score were most commonly used.

Two reviewers independently screened studies for inclusion; discrepancies were resolved through consensus.

Assessment of study quality
Two reviewers independently assessed the quality of included studies using the QUADAS tool, which included 14 criteria. Discrepancies between reviewers were resolved through consensus.

Data extraction
Two reviewers independently extracted or calculated true positives and negatives, and false positives and negatives to create a 2x2 table. The authors used their own software to calculate values. Discrepancies between reviewers were resolved through consensus.

Methods of synthesis
Sensitivity, specificity, positive and negative likelihood ratios and diagnostic odds ratios were pooled using standard meta-analysis methods to obtain summary results with their 95% confidence intervals. A bivariate generalised linear mixed model was also used to calculate pooled sensitivity and specificity. Pooled estimates were calculated for all reference standards and separately for each reference standard, and by the different cut-off values for the distress thermometer. Pooled estimates were also reported for the different cancer trajectories.

Summary receiver operator characteristics (SROC) curves were constructed (methods not reported) to assess the relationship between sensitivity and specificity for the distress thermometer versus all reference standards by different cut-off values. Area under the curve (AUC) values were also calculated; values from 0.9 to 1 indicating excellent diagnostic accuracy, 0.8 to 0.9 good accuracy, 0.7 to 0.8 fair, 0.6 to 0.7 poor, and 0.5 to 0.6 failed.

Statistical heterogeneity was assessed using \( I^2 \). Publication bias was assessed using Begg's and Egger's funnel plots and test.
Results of the review
Forty-two studies (14,808 patients, range 43 to 1,392) were included in the review. Six studies met all 14 quality assessment criteria; the remaining studies met either 12 or 13 criteria. Quality issues were apparent in terms of the reference standard being likely to correctly classify the target condition, and explanations for patient withdrawals. The authors considered all included to studies to have high quality.

Pooled estimates that compared the distress thermometer to all reference standards combined indicated that the best balance between sensitivity (0.81, 95% CI 0.79 to 0.82) and specificity (0.72, 95% CI 0.71 to 0.72) was at a cut-off score of 4 out of 10 on the thermometer scale (64 studies, table 2). The area under the curve (AUC) indicated good overall accuracy for the distress thermometer in detecting distress (AUC 0.83).

Pooled estimates for the different reference standards indicated that, when the distress thermometer was compared to the Hospital and Anxiety Depression Scale - Total at a cut-off value of 4, the balance between sensitivity (0.82) and specificity (0.73) was maximised (AUC 0.84).

Other results were fully reported in the review, including results for all cut-off values by the phase of the cancer trajectory (active treatment, survivorship, and end-of-life).

Authors’ conclusions
The distress thermometer was a good tool for screening cancer-related distress. When compared with commonly used reference standards, a cut-off score/value of 4 was considered optimal.

CRD commentary
The review question and supporting inclusion criteria were broadly stated. The literature search was limited to two databases and by search dates. The search did not appear to include attempts to locate unpublished data, which meant that potentially relevant data may have been missed. Formal assessment suggested there was some evidence of publication bias. Each stage of the review process was performed in duplicate, which reduced the potential for reviewer error and bias.

Study quality was assessed using appropriate criteria and the results were fully reported in the review. The authors suggested that the included studies were of high quality, but the poor ability of the reference standards to correctly classify target conditions questioned the reliability of the findings.

The studies were pooled using appropriate meta-analysis methods, but the patient and study characteristics reported indicated both patient and methodological differences across the studies.

The number of studies used to calculate diagnostic statistics for different reference standards and cut-off values was unclear. For example, the pooled estimates at a cut-off value of 4 was reported to have been calculated from 64 studies (table 2, column N1a). It was not clear which studies this figure referred to as only 42 studies were eligible for inclusion or whether some of the 42 included studies had been double counted.

This review included a large number of studies and patients. However, given the lack of a consistent and reliable reference standard, the authors’ conclusions should be considered carefully as the accuracy may be overestimated.

Implications of the review for practice and research
Practice: The authors stated that the evidence confirmed that a cut-off score of 4 was optimal for the distress thermometer. However, judgements on the appropriate care for patients should also include the skills of the doctors or health carers; the distress thermometer should not be used as a stand-alone diagnostic test.

Research: The authors stated that more studies were needed to further evaluate the performance of the distress thermometer in patients from different countries and with different types of cancer.

Funding
None.

Bibliographic details

PubMedID
24510195

DOI
10.1007/s00520-014-2143-1

Indexing Status
Subject indexing assigned by NLM

MeSH
Female; Humans; Likelihood Functions; Male; Mass Screening /methods; Middle Aged; Neoplasms /diagnosis /psychology; Odds Ratio; Psychometrics /methods; ROC Curve; Sensitivity and Specificity; Stress, Psychological /diagnosis

AccessionNumber
12014012823

Date bibliographic record published
27/02/2014

Date abstract record published
08/07/2014

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.