

UNIVERSIDAD PERUANA UNIÓN

ESCUELA DE POSGRADO

Unidad de Posgrado de Psicología



**Revisión sistemática de las propiedades psicométricas y precisión diagnóstica
del termómetro del estrés en adultos con cáncer**

**Trabajo Académico para obtener el Título de Segunda Especialidad Profesional de
Psicología Clínica y de la Salud**

Autores:

**Loida Esenarro-Valencia
Hernán Bernedo Del Carpio**

Asesor:

Nikolai Rodas Vera

Lima, enero de 2026

Resumen

Introducción: El distrés en cáncer es frecuente y su detección temprana permite una atención integral y oportuna. El Termómetro del Distrés (TD) es un instrumento ultrabreve de tamizaje, pero con variabilidad en métricas y puntos de corte según el contexto y el estándar de referencia. **Objetivo:** Sintetizar las propiedades psicométricas y de precisión diagnóstica del TD en adultos con cáncer, además de describir puntos de corte y estimar su desempeño global mediante el índice de Youden. **Métodos:** Se realizó una revisión sistemática según los lineamientos de PRISMA. Se realizaron búsquedas entre 2015 y julio de 2025 en PubMed/MEDLINE, PsycINFO, Cochrane Library, SciELO y LILACS (idiomas: español, inglés y portugués). Dos revisores realizaron la selección de estudios, la extracción de datos y la evaluación de calidad mediante COSMIN y QUADAS-2. Se calcularon métricas descriptivas y el índice de Youden para la evaluación del desempeño global. **Resultados:** Se incluyeron 27 estudios ($n = 51-768$; mediana = 247). La evidencia psicométrica se concentró (de tipo convergente), con evaluación limitada de confiabilidad. La precisión diagnóstica se reportó mediante el AUC en la mayoría de los estudios (mediana = .80, rango .60 – .88), con sensibilidad mediana de .83 (rango .50 a 1) y especificidad mediana de .68 (rango .36 a .97). El punto de corte ≥ 4 fue el más frecuente (45.8%), aunque se observaron umbrales más altos en contextos clínicos específicos. El índice de Youden mostró desempeño moderado (mediana .50, rango .21–.76), sugiriendo variación por contexto clínico y criterios de referencia. **Conclusiones:** El TD es útil como tamizaje en oncología adulta, pero la heterogeneidad de los umbrales, los patrones de referencia y la calidad del reporte limitan la estandarización de un punto de corte único. Se recomiendan estudios con mayor rigor metodológico y reporte de indicadores diagnósticos.

Palabras clave: neoplasias, termómetro del distrés, propiedades psicométricas, precisión diagnóstica, índice de Youden

Summary

Introduction: Cancer-related distress is common, and early detection enables comprehensive, timely care. The Distress Thermometer (DT) is an ultra-brief screening tool, though its metrics and cut-off points vary depending on the context and reference standard. **Objective:** To synthesize the psychometric properties and diagnostic accuracy of the DT in adults with cancer, describe cut-off points, and estimate overall performance using the Youden index. **Methods:** A systematic review was conducted following PRISMA guidelines. Searches were performed covering the period from 2015 to July 2025 across PubMed/MEDLINE, PsycINFO, the Cochrane Library, SciELO, and LILACS (languages: Spanish, English, and Portuguese). Two reviewers handled study selection, data extraction, and quality assessment using COSMIN and QUADAS-2. Descriptive metrics and the Youden index were calculated to evaluate overall performance. **Results:** Twenty-seven studies were included ($n = 51\text{--}768$; median = 247). Psychometric evidence focused primarily on convergent validity, with limited assessment of reliability. Diagnostic accuracy was reported via AUC in most studies (median = 0.80; range: 0.60–0.88), with a median sensitivity of 0.83 (range: 0.50–1.00) and a median specificity of 0.68 (range: 0.36–0.97). A cut-off point of ≥ 4 was the most common (45.8%), although higher thresholds were observed in specific clinical settings. The Youden index indicated moderate performance (median = 0.50; range: 0.21–0.76), suggesting variation based on clinical context and reference criteria. **Conclusions:** The DT is a useful screening tool in adult oncology; however, heterogeneity regarding thresholds, reference standards, and reporting quality limits the standardization of a single cut-off point. Studies with greater methodological rigor and comprehensive reporting of diagnostic indicators are recommended.

Keywords: neoplasms, distress thermometer, psychometric properties, diagnostic accuracy, Youden index

1. Introducción

En 2020 se estimaron más de 20 millones de casos nuevos de cáncer en el mundo y 9.7 millones de muertes por cáncer, con un impacto significativo en la población. Quienes padecen esta enfermedad pueden desarrollar malestar emocional que requiere atención psicológica desde el proceso diagnóstico, debido a la incertidumbre, el temor, el estigma asociado a la enfermedad y la necesidad de recibir apoyo psicológico oportuno (Donovan et al., 2014; Organización Mundial de la Salud, 2024). La literatura internacional reporta prevalencias con rangos amplios que van del 30 % al 80 %, dentro de los cuales algunos pacientes podrían cumplir criterios de trastornos mentales como el trastorno adaptativo, la depresión o la ansiedad. En etapas avanzadas, el distrés emocional llega a afectar entre el 30 % y el 70 % de los pacientes (Sánchez-Román et al., 2020). Su prevalencia varía según las condiciones clínicas, económicas o sociodemográficas (Bergerot et al., 2024).

Desde el modelo biopsicosocial, el distrés no se reduce únicamente a síntomas emocionales ni puede explicarse solo por variables biomédicas. Engel (1977) planteó que la enfermedad y el sufrimiento son producto de la interacción de múltiples dimensiones (biológicas, psicológicas y sociales), las cuales merecen ser atendidas en las áreas clínicas. En oncología, el distrés es considerado el sexto signo vital, lo que justifica su evaluación y atención clínica (Holland & Bultz, 2007).

La National Comprehensive Cancer Network (NCCN, 2026) define al distrés psicológico o malestar emocional como una experiencia desagradable que implica diferentes componentes psicológicos, emocionales, físicos, sociofamiliares o espirituales, que impactan en la persona a nivel de pensamiento, sentimientos y conducta

de tal manera que interfieren en la comprensión y afrontamiento de sus síntomas de enfermedad o tratamiento.

Detectar el distrés de forma rápida y eficaz es clave, ya que permite mejorar la adherencia al tratamiento, facilita intervenciones oportunas y contribuye al bienestar general del paciente. En este sentido, el Termómetro del Distrés (TD), desarrollado por la NCCN, surge como una herramienta breve, visual y de fácil uso, ideal como instrumento de primera línea para la evaluación inicial (Anexo A).

La utilidad del TD se respalda en diferentes contextos internacionales, Thalén-Lindström et al. (2013), en Suecia, reportaron una sensibilidad del 87 % y una especificidad del 73 %, con un análisis Receiver Operating Characteristic (ROC) y un área bajo la curva (AUC) de .86, utilizando como criterio de referencia la Escala Hospitalaria de Ansiedad y Depresión (HADS). Los resultados se mantuvieron consistentes a los 1, 3 y 6 meses. Lim et al. (2014), en Singapur, consideraron un punto de corte de 5, con una sensibilidad del 88% y especificidad del 81%, valores que se asociaron con el deterioro de la calidad de vida en diferentes tipos de cáncer. Donovan et al. (2014) reafirmaron este punto de corte de 4, identificado en el 52 % de los 57 artículos revisados y en una consulta realizada a 28 países pertenecientes a la Sociedad Internacional de Psicooncología (IPOS). Priede et al. (2014) observaron que el punto de corte óptimo de 4 mejora su rendimiento diagnóstico cuando se combina con entrevista clínica y el HADS. Asimismo, Graham-Wisener et al. (2021) reportaron que, en entornos de cuidados paliativos o enfermedad avanzada, el TD mostró mejor precisión con un corte de ≥ 5 o 6, especialmente para identificar ansiedad o depresión por separado.

Aunque el TD se utiliza ampliamente, persiste heterogeneidad en el reporte de los

puntos de corte, estándares de referencia y calidad metodológica. Esta variabilidad limita su interpretación clínica y justifica la realización de una síntesis sistemática que contemple sus propiedades psicométricas y estimadores de precisión diagnóstica.

El estudio se justifica por la necesidad de integrar la evidencia que actualmente es heterogénea sobre la utilidad del TD como indicador transdiagnóstico de malestar frente a la depresión o ansiedad. Desde una perspectiva práctica, la síntesis de sus propiedades psicométricas, estimadores de precisión diagnóstica y puntos de corte orientará decisiones de tamizaje y derivación, reduciendo la arbitrariedad en la elección de umbrales y favoreciendo una detección oportuna del malestar emocional en servicios. Asimismo, metodológicamente, la integración de la evaluación de riesgo de sesgo permitirá valorar la consistencia y calidad de la evidencia disponible.

El objetivo del estudio fue sintetizar las propiedades psicométricas y la precisión diagnóstica del TD en adultos con cáncer, describir el punto de corte óptimo y aportar una valoración crítica de la calidad y consistencia de la evidencia disponible. De manera específica, se buscó caracterizar los estudios distinguiendo su procedencia y contexto clínico, incluyendo el tipo de cáncer, estadio clínico y tratamiento; sintetizar los indicadores de sus propiedades psicométricas como validez y confiabilidad; sintetizar los indicadores de precisión diagnóstica incluyendo sensibilidad, especificidad y AUC; describir estimadores del desempeño global de los puntos de corte óptimos del TD y evaluar la calidad metodológica y el riesgo de sesgo de los estudios incluidos utilizando herramientas como el Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) para estudios psicométricos y el Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) para estudios diagnósticos,

respectivamente (Mokkink et al., 2018; Whiting et al., 2011).

2. Metodología

Diseño del estudio

Se realizó una revisión sistemática orientada a sintetizar evidencia empírica sobre las propiedades psicométricas y la precisión diagnóstica del TD en adultos con cáncer. Se desarrolló conforme a las recomendaciones de la guía de Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 (Page et al., 2021). Asimismo, se consideraron los lineamientos del sistema de clasificación de diseños de investigación en psicología propuesto por Ato et al. (2013), ubicándolo como un estudio secundario basado en estudios empíricos. Para estimar el desempeño global de los puntos de corte óptimos reportados, se calculó el índice de Youden (J), el cual es una forma muy usada de resumir en un valor único el equilibrio entre sensibilidad y especificidad, así como la efectividad global del test para un umbral determinado (Youden, 1950).

El protocolo se registró en la International Prospective Register of Systematic Reviews (PROSPERO) con el código CRD420251176992, disponible en <https://www.crd.york.ac.uk/PROSPERO/view/CRD420251176992>.

Fuentes de información

Se realizaron búsquedas sistemáticas en las siguientes bases de datos electrónicas: PubMed/MEDLINE, PsycINFO, Cochrane Library como bases de datos relevantes en el área de salud, además de SciELO y LILACS de alcance regional. Para la elaboración de la estrategia de búsqueda se incluyeron los siguientes descriptores: “neoplasias”, “termómetro del distrés”, “validez”, “confiabilidad”, “sensibilidad”,

“especificidad”, “propiedades psicométricas”, “precisión diagnóstica” en población adulta; además de términos del Medical Subject Headings (MeSH) y Descriptores en Ciencias de la Salud (DeCS) organizados en una tabla (Anexo B) según la población, intervención y desenlace requeridos. En PubMed/MEDLINE se utilizaron ajustadores de campo para la búsqueda en título y resumen, además de filtros de temporalidad y edad; en PsylInfo, ajustadores de resumen y filtros por edad y temporalidad; en Cochrane Library, ajustadores de título, resumen y palabras clave; y en SciELO y LILACS, términos clave (Anexo C). La búsqueda se realizó el 24 de agosto de 2025 por los revisores y la información fue organizada en una base de datos de datos elaborada en Microsoft Excel.

Se incluyeron estudios publicados entre 2015 y 2025 en inglés, español y portugués, que evaluaran al menos una propiedad psicométrica del TD, como validez, confiabilidad o precisión diagnóstica. La población de interés estuvo conformada por personas adultas con diagnóstico confirmado de cáncer, sin restricción de tipo o estadio. Se consideraron estudios con diseño transversal o longitudinal que reportaran dichas medidas. Se excluyeron estudios realizados en población pediátrica y no oncológica, así como revisiones narrativas, revisiones sistemáticas, libros, cartas del editor, pósteres o resúmenes de congresos. Se excluyeron, también, estudios que no reportaran medidas de validez, confiabilidad o indicadores diagnósticos, como sensibilidad, especificidad, área bajo la curva u otros estimadores relacionados.

Estrategia de búsqueda:

La estrategia de búsqueda (Anexo C) fue desarrollada combinando descriptores controlados (MeSH/DeCS) y términos libres, a fin de identificar con precisión los estudios relevantes en diferentes bases de datos seleccionadas relacionadas con las áreas de

salud y psicológica.

Esta estrategia fue diseñada y validada cuidadosamente por dos revisores (LEV, HBDC), para adaptarse a la estructura y características de cada plataforma de búsqueda (Anexo C). Se emplearon operadores booleanos (AND, OR) para conectar los términos relacionados con el instrumento, los desenlaces psicométricos y la población de interés. Asimismo, se utilizaron ajustadores de campo (como título, resumen, términos MeSH/DeCS) y filtros según el año de publicación (2015–2025) y la edad de los participantes (adultos).

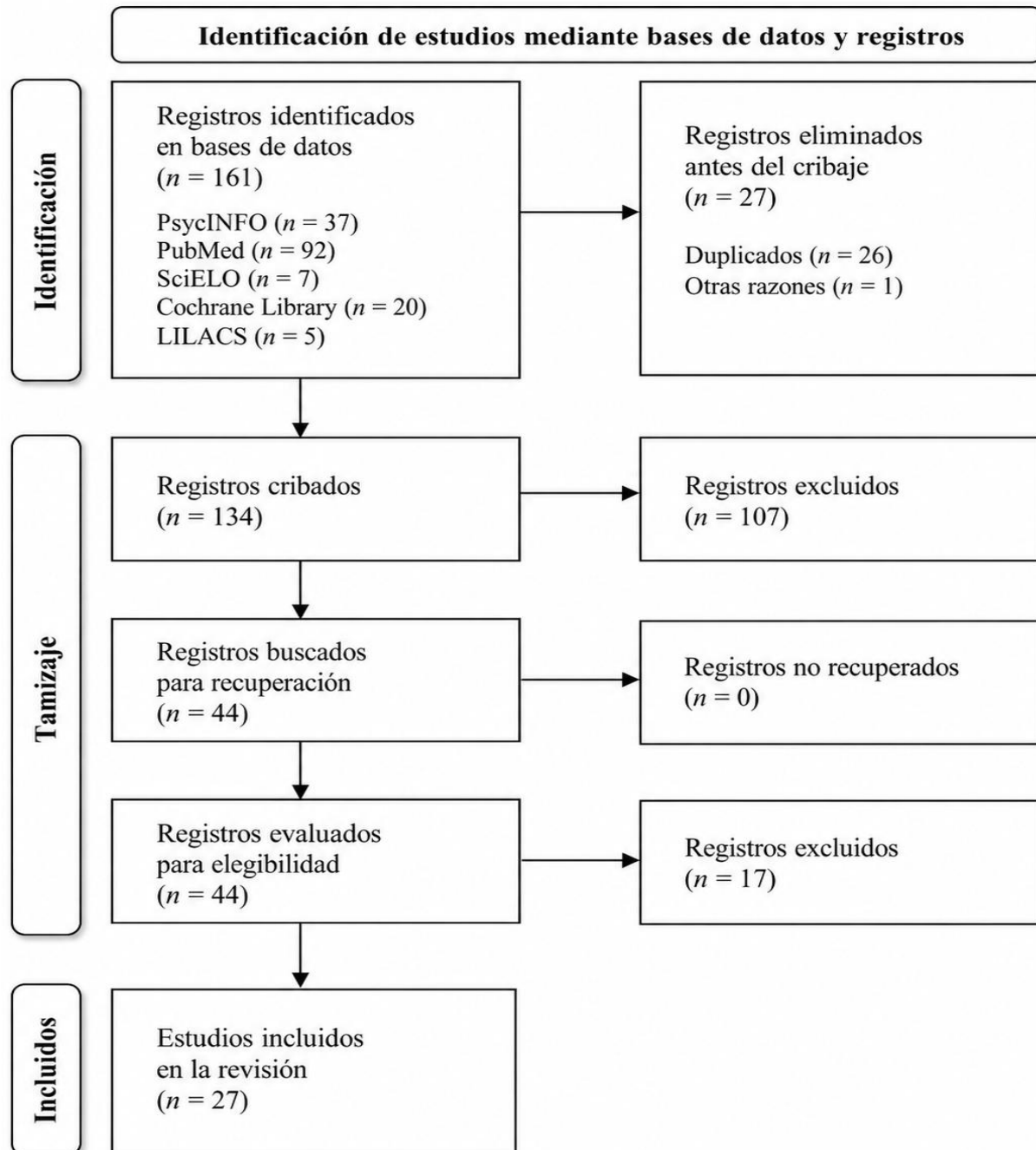
Extracción de datos

De manera independiente, los revisores extrajeron datos generales (autor, año, país de procedencia), muestra, diseño de investigación y contexto clínico (edad, diagnóstico, tipo y estadiaje de cáncer, tratamiento, modalidad de aplicación y evaluador); así como los indicadores o métricas relacionadas con validez, confiabilidad, precisión diagnóstica y puntos de corte. Las discrepancias fueron resueltas por consenso. La información se almacenó en una matriz de Excel. Asimismo, se utilizó la plataforma Rayyan (<https://www.rayyan.ai/>) para la gestión de datos, el proceso de tamizaje y la evaluación de los estudios incluidos, así como para establecer la estructura preliminar de la tabla de extracción de datos.

El siguiente diagrama de flujo (https://estech.shinyapps.io/prisma_flowdiagram/), basado en la declaración PRISMA (Figura 1), sintetiza el proceso de búsqueda, selección y los estudios incluidos para el análisis. Asimismo, se utilizaron herramientas digitales para optimizar la visualización de gráficos y optimizar la visualización de los datos, sin intervención en el análisis o interpretación de estos (OpenAI, 2026).

Figura 1.

Diagrama de flujo del proceso de selección de estudios sobre el termómetro del distrés en adultos con cáncer.



Nota: Adaptado de Page et al. (2021).

Evaluación de riesgo de sesgo

Se aplicó la lista de chequeo COSMIN (<https://www.cosmin.nl/>) para evaluar el riesgo de sesgo metodológico de los estudios incluidos. Siguiendo la recomendación del manual COSMIN, se seleccionaron las cajas (boxes) correspondientes a las propiedades de medición pertinentes, siendo la Box 9 (prueba de hipótesis para la validez de constructo) así como el Box 6 (fiabilidad). Para los estudios de precisión diagnóstica se utilizó el QUADAS-2, el cual valora el riesgo de sesgo en cuatro dominios: selección de pacientes, prueba índice, estándar de referencia y flujo y tiempo. La evaluación fue realizada de manera independiente por los autores (LEV-HBDC), y las discrepancias se resolvieron mediante debate y discusión (Mokkink et al., 2018; Whiting et al., 2011).

Síntesis de datos

Los resultados se sintetizaron de manera descriptiva, considerando características de los estudios incluidos, las propiedades psicométricas, indicadores de precisión y puntos de corte. Se calculó el índice de Youden a partir de los estudios que reportaron sensibilidad y especificidad para estimar el desempeño global del punto de corte informado (Youden, 1950).

3. Resultados

Se identificaron 161 registros en las bases de datos de PubMed, PsycINFO, SciELO, LILACS y Cochrane Library (Figura 1). Luego de la eliminación de 26 duplicados y uno por presentar problemas de ilegibilidad, quedaron 134 artículos. Posteriormente, se realizó la evaluación de título y *abstract*, y como resultado del screening se obtuvieron 44 artículos para valorar según criterios de inclusión/exclusión de manera crítica. Finalmente, se excluyeron cinco registros por utilizar un instrumento diferente; tres por

pertenecer a una población diferente, siete estudios por tener desenlaces diferentes, uno por límite de tiempo, y uno por tipo de publicación, quedando 27 artículos incluidos para la síntesis cualitativa.

Los estudios incluyeron tamaños muestrales heterogéneos, entre 51 y 768 participantes, con una mediana de 247. Los estudios proceden de regiones geográficas de Europa (14), Asia (7), América (4) y Oceanía (2) y, de ellos, dos estudios corresponden a países de habla hispana (Anexo D).

El Anexo E resume las características de 27 estudios que evaluaron el rendimiento del TD con participantes entre 18 y 95 años. Se incluyeron tres estudios oncogeriátricos (Boudin et al., 2022; Ratnasekera et al., 2023; Zuber et al., 2025). Predominaron los diagnósticos mixtos (55.6%); específicamente, fueron más frecuentes el cáncer de mama (14.8%), el cáncer ginecológico y el colorrectal (7.4%), tumores cerebrales, tumores sólidos, cáncer de cabeza y cuello y cáncer oromaxilofacial (3.7% cada uno). Tres estudios (11.1%) abordaron el cáncer avanzado y en cuidados paliativos y reportaron también la presencia de sintomatología dolorosa (Ascencio et al., 2024; Guan et al., 2019; Van Lander et al., 2019).

Dos estudios reportan cáncer avanzado metastásico, mientras que 10 estudios (37%) reportan de manera variada estadios desde el I al IV. El tratamiento fue mayoritariamente mixto, es decir, incluyó cirugía, quimioterapia u otras modalidades terapéuticas y alcanzó un 33.3%, seguido del tratamiento prequirúrgico (11.1%), cirugía exclusiva (3.7%), quimioterapia exclusiva (11.1%), tratamiento del dolor (3.7%), y otros reportan tratamiento activo sin especificar o sin tratamiento.

Los evaluadores o aplicadores, en su mayoría conformaron los equipos

multidisciplinarios (11.1%) habituales en oncología como médicos (3.7%) o residentes de medicina (Alosaimi et al., 2018), personal de enfermería (14.8%) (Andreu et al., 2019; Civilotti et al., 2020; Kyranou et al., 2020; Linehan et al., 2017); y finalmente psicología con un 3.7%. En su mayoría, la aplicación se realizó mediante autorreporte, a través de formatos de lápiz y papel y con una menor proporción mediante modalidad en línea (Linehan et al., 2017; Olesen et al., 2018) o no reportado (Ascencio et al., 2024).

Los criterios de referencia fueron heterogéneos, predominando la Escala Hospitalaria de Ansiedad y Depresión (HADS), además del Inventario de Síntomas Abreviado (BIS-18) y medidas específicas para depresión, ansiedad o estrés como Cuestionario de Salud del Paciente de 9 ítems (PHQ-9), Cuestionario de salud de 12 ítems de China (CHQ-12), la Escala de Ansiedad, depresión y estrés de 21 ítems (DASS-21), la escala de distrés psicológico de Kessler (K10) el Sistema de Evaluación de Síntomas de Edmonton Revisado (ESAS-r), la Escala de Ansiedad Generalizada (GAD-7), la Escala de depresión geriátrica de Yesavage (GDS), Escala de funcionalidad emocional; así como instrumentos de calidad de vida y entrevistas o criterios diagnósticos como el Structured Clinical Interview for DSM Disorders (SCID) del Manual diagnóstico y estadístico de los trastornos mentales, quinta edición (DSM-5), la Mini International Neuropsychiatric Interview (MINI) de la Clasificación Internacional de Enfermedades, décima revisión (CIE-10) o criterios clínicos.

La valoración del riesgo de sesgo de los estudios incluidos se muestra en el Anexo F. Usando el QUADAS-2, las limitaciones centrales se relacionan con la selección de pacientes, los puntos de corte derivados de la misma muestra, los criterios o pruebas de referencia y, finalmente, en el dominio de flujo y tiempo, presentaron bajo riesgo de

sesgo. En general, los resultados sugieren una calidad metodológica variable, es decir, un potencial sesgo de sobreestimación de la precisión diagnóstica.

En COSMIN (Anexo G) la evidencia se concentró en la validez de constructo a través del análisis de validez convergente. De modo general, la calidad metodológica fue adecuada, con limitaciones derivadas del uso de instrumentos de autorreporte como comparadores, lo que introduce sesgo y una medición indirecta. En los estudios de confiabilidad, la calidad fue adecuada, y en contraste con la validez de constructo la evidencia es escasa.

La evidencia psicométrica (Tabla 1) se centró en la validez de constructo, principalmente con análisis de validez convergente, mientras que la confiabilidad, fue reportada con menor frecuencia.

En los estudios de precisión diagnóstica (Tabla 2), el AUC fue reportado en 24 estudios (88.9%), con una mediana de .80 y un rango de .60 a .88. La sensibilidad presentó una mediana de .83 (rango: de .50 a 1.0); por otro lado, la especificidad presentó una mediana de .68 (rango .36 a .97). En torno a los puntos de corte, estos se distribuyen en valores heterogéneos, y casi en su mayoría (45.8%) corresponden a valores mayores o iguales a 4. Puntos de corte mayores o iguales a 5 y 7 han sido descritos en 8 estudios (33.3%), mientras que los puntajes menores o iguales a 2 y 3 fueron reportados en menor proporción (12.5%). Un solo estudio reportó un punto de corte de 4.5.

El índice de Youden (Anexo H) estimado como medida de desempeño global en 24 contrastes diagnósticos, mostró un desempeño global moderado con una media de .48, mediana $J = .50$, y rango de .21 a .76. Conev et al. (2019) no fue incluido en este

cálculo por corresponder a una medida repetida y no a un estándar de referencia. Al agrupar por umbrales (Tabla 3), ≥ 4 fue el punto de corte más frecuente (mediana $J = .52$) frente a ≥ 5 (mediana $J = .47$). Existe variabilidad, lo que sugiere diferencias relevantes según el contexto, los criterios de comparación y el criterio de referencia utilizado.

Tabla 1.

Propiedades psicométricas del Termómetro del Distrés en los estudios incluidos.

Autor (año)	Propiedad psicométrica	Instrumento	Evidencia (contextualizada)	Indicador
Alosaimi et al. (2018)	Validez de constructo (convergente)	HADS	El TD mostró asociación moderada con síntomas ansioso-depresivos, consistente con convergencia esperada.	$r = .50, p < .001$
Andreu et al. (2019)	Validez de constructo (convergente)	BSI-18	El TD se relacionó moderada-alta con malestar psicológico global y dominios afectivos.	$r = .67, p < .001$
Antoniadis et al. (2015)	Validez de constructo (convergente)	HADS	Se evidenció convergencia moderada con síntomas ansioso-depresivos.	$r = .39, p < .001$
Antoniadis et al. (2015)	Fiabilidad (consistencia interna)	TD	La consistencia interna fue adecuada para la medición del constructo.	$\alpha = .80$
Bergerot et al. (2024)	Validez de constructo (convergente)	SupportScreen	Alta convergencia con un instrumento biopsicosocial multidimensional.	$r = .80$, IC 95% [.75, .84]
Linehan et al. (2017)	Validez de constructo (convergente)	DASS-21	Asociación moderada-alta con depresión, ansiedad y estrés.	$r = .45-.64, p < .001$
Nguyen et al. (2021)	Validez de constructo (convergente)	PHQ-9	Asociación moderada con síntomas depresivos.	$r = .53$, IC 95% [.44, .61]
Nguyen et al. (2021)	Validez de constructo (convergente)	GAD-7	Asociación moderada con síntomas de ansiedad.	$r = .51$, IC 95% [.42, .59]
Olesen et al. (2018)	Validez de constructo (ROC)	HADS	Capacidad discriminativa aceptable en comparación con HADS.	AUC = .73, IC 95% [.64, .81]
Ratnasekera et al. (2023)	Fiabilidad (acuerdo)	Clasificación DSM-IV	Acuerdo en la clasificación diagnóstica.	$\kappa = .89, p < .001$
Robbeson et al. (2019)	Validez de constructo (convergente)	SF-36	Asociación inversa con calidad de vida, consistente con el constructo.	$r = -.45, p < .01$
Sarenmalm et al. (2018) *	Validez de constructo (convergente)	QLQ-C30	Asociación fuerte con dimensión emocional e inversa con calidad de vida global.	$r = .74$; $r = -.57, p < .001$
Zheng et al. (2019)	Validez de constructo (convergente)	HADS	Alta convergencia con síntomas ansioso-depresivos.	$r = .78, p < .05$
Van Lander et al. (2019)	Fiabilidad (acuerdo)	Criterio clínico	Concordancia moderada entre evaluación clínica y autorreporte.	$\kappa \approx .52$
Zuber et al. (2025)	Validez de constructo (convergente)	GDS	Asociación moderada con depresión geriátrica.	$r = .41, p < .001$

Nota: TD = Termómetro del Distrés; IC = intervalo de confianza; AUC = área bajo la curva; κ = coeficiente kappa. Los coeficientes de correlación (r) y valores de AUC se presentan sin cero inicial, conforme a APA 7. Nguyen et al. (2021) aporta dos análisis independientes (PHQ-9 y GAD-7).

Tabla 2*Precisión diagnóstica del Termómetro del Distrés en los estudios incluidos.*

Autor (año)	Sensibilidad	Especificidad	AUC (IC 95%)	Punto de corte	Criterio de referencia
Alosaimi et al. (2018)	.70	.63	.76 (.73–.88)	≥4	HADS
Andreu et al. (2019)	.89	.59	.81 (.74–.87)	≥5	BSI-18
Antoniadis et al. (2015)	.73	.79	.80 (.67–.94)	≥5	HADS
Ascencio et al. 8/07/2026 18:00:00	.76	.54	.69 (NR)	≥5	ESAS-r
Boudin et al. (2024; 2022)	.50	.97	.67 (.40–.95)	≥4	DSM-5
Chiou et al. (2017)	.72	.80	.79	≥4	CHQ-12
Civilotti et al. (2020)	.89	.64	.85 (.81–.89)	≥4	HADS
Conev et al. (2019)	.66	.53	.60 (.53–.67)	≥4	Medida repetida
Ghazali et al. (2017)	.84	.75	.88	≥4	UWQOL
Guan et al. (2019)	.86	.53	.75 (.70–.81)	≥5	HADS
Kuchen et al. (2025)	.81	.46	.70 (.66–.74)	≥5	EF
Kyranou et al. (2020)	.85	.69	.79	≥4	MINI (CIE-10)
Liu et al. (2024)	1.00	.72	.86 (.86–.93)	≥4	HADS
Martinez et al. (2015)	.93	.57	.84 (.80–.87)	≥5	BSI-18
Nguyen et al. (2021), PHQ-9	.87	.58	.81	≥4	PHQ-9
Nguyen et al. (2021), GAD-7	.92	.58	.82	≥4	GAD-7
Olesen et al. (2018)	.93	.36	.73 (.64–.81)	≥2	HADS
Ploos Van Amstel et al. (2017)	.73	.84	.60 (.53–.67)	≥7	HADS
Ratnasekera et al. (2023)	.98	.78	.84	≥4	DSM-IV
Recklitis et al. (2016)	.68	.78	.80	≥5	DSM-IV (SCID)
Robbeson et al. (2019)	.53	.68	.66 (.54–.78)	≥4.5	SF-36
Seib et al., (2024)	.92	.73	.88 (.85–.92)	≥5	K-10
Van Lander et al., (2019)	.83	.71	NR	≥3	Criterio clínico
Zheng et al. (2019)	.78	.79	.80	≥3	HADS
Zuber et al. (2025)	.66	.68	.70	≥4	GDS

Tabla 3*Desempeño global del TD según el índice de Youden por punto de corte*

Punto de corte	Contraste diagnóstico	Mediana de J	Rango de J
≥2	1	.29	.29
≥3	2	.56	.54 – .57
≥4	11	.52	.33 -.76
≥4.5	1	.21	.21
≥5	8	.47	.27 -.65
≥7	1	.57	.57

Nota. TD = Termómetro del Distrés; J = índice de Youden. Los valores de J se calcularon mediante la fórmula sensibilidad + especificidad – 1. Conev et al. fue excluido del cálculo por no contar con un estándar de referencia independiente.

4. Discusión

El Termómetro del Distrés (TD) ha mostrado su utilidad como herramienta de tamizaje ultrabreve en oncología adulta (McElroy et al., 2022). Ha sido recomendado por las guías de manejo del distrés de la NCCN para su uso rutinario en el abordaje del malestar emocional (Ownby, 2019). El distrés ha sido propuesto como el sexto signo vital y el TD como herramienta para su valoración y monitoreo rutinario en diferentes contextos clínicos (Howell & Olsen, 2011). La presente revisión incluyó 27 estudios, en los cuales el TD mostró un rendimiento de cribado aceptable, con un punto de corte de ≥4 como umbral de tamizaje. Esto es consistente con otras estimaciones que reportaron un rendimiento combinado de sensibilidad .81, especificidad .72 y AUC .83, identificando ≥4 como punto de corte con mejor desempeño frente a distintos criterios de referencia y mostrando que dicho umbral maximizó el equilibrio entre sensibilidad y especificidad. En este sentido, los resultados tienden a reflejar el patrón de sensibilidad y especificidad, aunque fueron ligeramente menores al comparar la mediana del AUC (.80) con el valor

previamente reportado (.83). Asimismo, el punto de corte ≥ 4 representó una proporción significativa (45.8 % de los estudios que reportaron dicho umbral), seguido de puntos de corte mayores (≥ 5 y ≥ 7 , respectivamente), lo que implica la necesidad de observar el contexto (fase clínica, tratamiento, entre otras condiciones) para la toma de decisiones oportunas en la atención (Ma et al., 2014).

Al integrar los datos del rendimiento diagnóstico y la evaluación del riesgo de sesgo, se observaron valores más elevados en sensibilidad y especificidad en estudios con limitaciones metodológicas, mientras que aquellos con menor riesgo de sesgo reportaron estimaciones más moderadas (Kuchen et al., 2025; Recklitis et al., 2016). Con relación a los puntos de corte, al ser agrupados, se aprecia que el umbral de valores mayores o iguales a 4 permite optimizar la sensibilidad (favoreciendo la detección de casos). Los valores de 5 y 7, referidos anteriormente, sugieren que el rendimiento del TD es dependiente del contexto, dado que incrementan la especificidad y reducen la captación de falsos positivos. En la misma línea, los valores del índice de Youden sugieren que los valores mayores o iguales a 4 ofrecen la posibilidad de mejorar el equilibrio entre sensibilidad y especificidad, orientando la selección de umbrales según las necesidades particulares

Varios de los estudios proceden de ámbitos hospitalarios y de contextos particulares, como los cuidados paliativos o la medicina del dolor, y debido a ello, se recomienda utilizar puntos de corte más altos para ganar especificidad. En población con cáncer avanzado, Graham-Wisener et al. (2021); hallaron que valores mayores o iguales a 6 fueron óptimos para identificar el distrés total, y los mayores o iguales a 4, lo fueron para depresión. En general se han recomendado valores mayores a 5 cuando se prioriza

sensibilidad, aunque esto implique un incremento de falsos positivos (Ascencio et al., 2024). Esto puede ocurrir cuando la carga de malestar es alta, como en el caso de cuidados paliativos o un diagnóstico reciente (Kuchen et al., 2025). Desde la perspectiva del impacto en los servicios, el uso diferencial del TD permite obtener un umbral de riesgo que promueva la respuesta temprana del profesional o equipo de trabajo. Esto podría contribuir a la reducción de costos a mediano o largo plazo y ser coherente con lo postulado sobre la detección del sexto signo vital y enfoques de atención oncológica, de rehabilitación y paliativa escalonada (Carlson & Bultz, 2004; Jefford et al., 2022).

Aunque la mayoría de los estudios sobre el TD han utilizado análisis ROC y valores de sensibilidad y especificidad para definir puntos de corte, pocos han incorporado explícitamente métricas integradoras como el índice de Youden para optimizar la selección del umbral (Mitchell, 2007). Por ello, esta revisión constituye una aproximación más sistemática para la interpretación del rendimiento del TD. Tanto en los estudios incluidos como en la literatura previa persiste heterogeneidad en los criterios de referencia (HADS, BSI-18, CHQ-12, entre otros), puntos de corte y frecuentes carencias en el reporte de tablas 2×2 , intervalos de confianza del 95 % o procedimientos de cegamiento. Esta variabilidad metodológica limita la comparabilidad y genera la posibilidad de sesgo y, en consecuencia, afecta la precisión de las estimaciones.

Además, se advierte que el HADS, uno de los instrumentos más utilizados en el ámbito clínico y de investigación, presenta aspectos relacionados con la validez de constructo y puede capturar un factor general de distrés, ya que puede utilizarse la versión orientada a depresión o la versión general que incluye ansiedad. Esto afecta la comparabilidad y especificidad del TD cuando se utiliza como estándar de referencia

(Coyne & van Sonderen, 2012; Graham-Wisener et al., 2021). En consecuencia, la interpretación de la AUC, así como de la sensibilidad y especificidad, debe considerar el estándar de referencia empleado, el cual suele estar escasamente reportado, además del umbral específico y la calidad del reporte metodológico.

5. Conclusiones

La evidencia disponible resalta que el TD es un instrumento breve, aceptable y clínicamente útil para el tamizaje o cribado del distrés o malestar emocional en oncología adulta. Su rendimiento diagnóstico, estimado por el AUC, se ubica en rangos compatibles con utilidad clínica, lo que favorece su implementación en el entorno clínico, consolidándose como una herramienta con umbrales operativos útiles para la detección inicial del distrés. Aun así, este rendimiento debe interpretarse como suficiente para el tamizaje, pero no como una herramienta diagnóstica, sino como un instrumento inicial de identificación de casos que requiere de una evaluación complementaria, considerándose dentro de modelos de atención escalonada.

Los puntos de corte tienden a agruparse en torno al rango de ≥ 4 a ≥ 5 , manteniendo el punto de corte ≥ 4 como el más utilizado para optimizar el tamizaje, mientras que puntos mayores son apropiados en contextos de mayor complejidad, como áreas hospitalarias, evaluaciones prequirúrgicas, diagnóstico reciente, enfermedad avanzada y cuidados paliativos. En ese sentido, se reafirma la utilidad del TD con un perfil dependiente del contexto clínico y asistencial.

Desde la perspectiva psicométrica, la evidencia disponible se concentra de manera predominante en la validez de constructo, principalmente la validez convergente. En contraste, la confiabilidad ha sido escasamente evaluada, limitándose principalmente

a la interpretación del desempeño del TD en comparación con instrumentos afines, lo que evidencia vacíos relevantes en la evaluación de la estabilidad y consistencia del instrumento. La incorporación del índice de Youden permite una aproximación más integrada del rendimiento del TD a partir del análisis de la sensibilidad y especificidad, orientando la selección de umbrales según necesidades particulares.

Entre las limitaciones de esta revisión, se reconoce que no se incluyeron otras bases de datos de cobertura interdisciplinaria, lo que podría haber introducido un sesgo en la recopilación o localización de estudios. Asimismo, la diversidad metodológica entre estudios limitó la comparabilidad directa y redujo la posibilidad de una síntesis cuantitativa única, por lo que no se aplicaron procedimientos estadísticos propios del metaanálisis de exactitud diagnóstica ni de otros análisis psicométricos. Estos aspectos destacan la necesidad de estudios futuros con mayor rigor metodológico, uso de estándares clínicos y reporte completo de indicadores que permitan consolidar la evidencia sobre el uso del TD en distintos contextos.

1. Recomendaciones

Los resultados de esta revisión respaldan el uso del Termómetro del Distrés (TD) como un instrumento útil para el cribado del malestar emocional en adultos con cáncer. Su aplicación no debe basarse en un punto de corte fijo, sino en una selección de umbrales sensibles al contexto clínico, considerando las diferentes variables propias de la enfermedad, el ámbito asistencial y el objetivo de uso.

Las investigaciones futuras deben priorizar estudios de precisión diagnóstica con mayor rigor metodológico, incluyendo estándares de referencia clínicos bien estructurados, reporte completo de indicadores (tablas de contingencia 2×2 ,

sensibilidad, especificidad e intervalos de confianza) y el uso de métricas integradoras (como el índice de Youden), para optimizar la selección de puntos de corte. De esta manera, estudios de utilidad clínica podrán examinar el impacto de la derivación, la adherencia al tratamiento y la reducción del estrés de manera longitudinal. Asimismo, es necesario fortalecer la evidencia sobre sus propiedades psicométricas en modalidades digitales y remotas como telesalud.

Finalmente, se requiere mayor consistencia metodológica y transparencia en el reporte de los resultados, con el respaldo de lineamientos estandarizados (como COSMIN o QUADAS-2) para la evaluación del riesgo de sesgo, con el propósito de consolidar la evidencia disponible, favorecer su síntesis e interpretación en diferentes contextos clínicos.

Referencias

- Alosaimi, F. D., Abdel-Aziz, N., Alsaleh, K., AlSheikh, R., AlSheikh, R., & Abdel-Warith, A. (2018). Validity and feasibility of the Arabic version of distress thermometer for Saudi cancer patients. *PloS One*, 13(11), e0207364. <https://doi.org/10.1371/journal.pone.0207364>
- Andreu, Y., Lopez, P. M., & Garrido, M. J. G. (2019). Use of the Distress Thermometer in cancer survivors: Convergent validity and diagnostic accuracy in a Spanish sample. *Oncology Nursing Forum*, 46(4), 442-450. <https://doi.org/10.1188/19.ONF.442-450>
- Antoniadis D, Lavrentiadis G, Papadopoulos V.N., Ierodiakonou I, & Garyfallos G. (2015). The Validity of the Distress Thermometer in Greek Colon Cancer Patients. *Acta Chirurgica Belgica*, 115(6), 387-392. <https://doi.org/10.1080/00015458.2015.11681138>
- Ascencio, L., Allende, S. R., & Peña, A. (2024). Assessment of emotional distress in palliative care: Edmonton Symptom Assessment System-revised (ESAS-r) vs Distress Thermometer. *Palliative & Supportive Care*, 22(2), 258-264. (rayyan-303948805). <https://doi.org/10.1017/S1478951523001530>
- Ato, M., López, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29, 1038-1059.
- Bergerot, C. D., Razavi, M., Bergerot, P. G., De Domenico, E., Birelo Lopes, Clark, K. L., Loscalzo, M., Pal, S. K., & Dale, W. (2024). Validation of a biopsychosocial distress screening tool, supportscreen, in a Brazilian cancer center. *Psycho-Oncology*, 33(10), 1-9. <https://doi.org/10.1002/pon.70001>
- Boudin, G., Solem Laviec, H., Ghewy, L., Le Bon, P., Lebaube, S., Machavoine, J.-L., Denhaerynck, J., Morello, R., & Beauplet, B. (2022). A prospective observational cohort study to screen major depressive disorders in geriatric oncology-Comparison of different scales. *European Journal of Cancer Care*, 31(4), e13591. <https://doi.org/10.1111/ecc.13591>
- Carlson, L. E., & Bultz, B. D. (2004). Efficacy and medical cost offset of psychosocial interventions

- in cancer care: Making the case for economic analyses. *Psycho-Oncology*, 13(12), 837-849; discussion 850-856. <https://doi.org/10.1002/pon.832>
- Chiou, Y.-J., Lee, C.-Y., Li, S.-H., Chong, M.-Y., Lee, Y., & Wang, L.-J. (2017). Screening for psychologic distress in Taiwanese cancer inpatients using the national comprehensive cancer network distress thermometer: The effects of patients' sex and chemotherapy experience. *Psychosomatics: Journal of Consultation and Liaison Psychiatry*, 58(5), 496-505. <https://doi.org/10.1016/j.psych.2017.03.016>
- Civilotti, C., Acquadro, D., Santagata, F., Varetto, A., & Stanizzo, M. (2020). The use of the Distress Thermometer and the Hospital Anxiety and Depression Scale for screening of anxiety and depression in Italian women newly diagnosed with breast cancer. *Supportive Care in Cancer*, 28(10), 4997-5004. <https://doi.org/10.1007/s00520-020-05343-x>
- Conev, N., Donev, I., & Stoyanov, D. (2019). One-minute time interval estimation as a novel ultrashort tool for distress screening. *Supportive Care in Cancer*, 27(6), 2031-2037. <https://doi.org/10.1007/s00520-018-4461-1>
- Coyne, J. C., & van Sonderen, E. (2012). No further research needed: Abandoning the Hospital and Anxiety Depression Scale (HADS). *Journal of Psychosomatic Research*, 72(3), 173-174. <https://doi.org/10.1016/j.jpsychores.2011.12.003>
- Donovan, K. A., Grassi, L., McGinty, H. L., & Jacobsen, P. B. (2014). Validation of the distress thermometer worldwide: State of the science. *Psycho-Oncology*, 23(3), 241-250. <https://doi.org/10.1002/pon.3430>
- Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science (New York, N.Y.)*, 196(4286), 129-136. <https://doi.org/10.1126/science.847460>
- Ghazali N, Roe B, Lowe D, Tandon S, Jones T, Brown J, Shaw R, Risk J, & Rogers S.N. (2017). Screening for distress using the distress thermometer and the University of Washington Quality of Life in post-treatment head and neck cancer survivors. *European Archives of Oto-Rhino-Laryngology*, 274(5), 2253-2260. <https://doi.org/10.1007/s00405-017-4474-2>

- Graham-Wisener, L., Dempster, M., Sadler, A., McCann, L., & McCorry, N. K. (2021). Validation of the distress thermometer in patients with advanced cancer receiving specialist palliative care in a hospice setting. *Palliative Medicine*, 35(1), 120-129. <https://doi.org/10.1177/0269216320954339>
- Guan, B., Wang, K., Shao, Y., Cheng, X., Hao, J., Tian, C., Chen, L., Ji, K., & Liu, W. (2019). The use of distress thermometer in advanced cancer inpatients with pain. *Psycho-oncology*, 28(5), 1004-1010. <https://doi.org/10.1002/pon.5032>
- Holland, J. C., & Bultz, B. D. (2007). The NCCN guideline for distress management: A case for making distress the sixth vital sign. *Journal of the National Comprehensive Cancer Network*, 5(1), 3-7.
- Howell, D., & Olsen, K. (2011). Distress-the 6th vital sign. *Current Oncology (Toronto, Ont.)*, 18(5), 208-210. <https://doi.org/10.3747/co.v18i5.790>
- Jefford, M., Howell, D., Li, Q., Lisy, K., Maher, J., Alfano, C. M., Rynderman, M., & Emery, J. (2022). Improved models of care for cancer survivors. *Lancet (London, England)*, 399(10334), 1551-1560. [https://doi.org/10.1016/S0140-6736\(22\)00306-3](https://doi.org/10.1016/S0140-6736(22)00306-3)
- Kuchen, R., Singer, S., Schranz, M., Doerner, L., Rieger, D., Steinbach, J. P., Ronellenfitsch, M. W., Voss, M., Kessler, A. F., Nickl, V., Misch, M., Onken, J. S., Rapp, M., Nadji-Ohl, M., Mehlitz, M., Meixensberger, J., Fehrenbach, M. K., Keric, N., Ringel, F., ... Renovanz, M. (2025). Distress screening in patients with high-grade glioma: Diagnostic accuracy in relation to a structured clinical interview in a multicenter cluster-randomized controlled trial. *Supportive Care in Cancer*, 33(8), 737. <https://doi.org/10.1007/s00520-025-09810-1>
- Kyranou, M., Varvara, C., Papathanasiou, M., Diakogiannis, I., Zafeiropoulos, K., Apostolidis, M., Papandreou, C., & Syngelakis, M. (2020). Validation of the Greek version of the Distress Thermometer compared to the Clinical Interview for Depression. *BMC Psychiatry*, 20(Journal Article). <https://doi.org/10.1186/s12888-020-02926-0>
- Lim, H. A., Mahendran, R., Chua, J., Peh, C.-X., Lim, S.-E., & Kua, E.-H. (2014). The Distress

- Thermometer as an ultra-short screening tool: A first validation study for mixed-cancer outpatients in Singapore. *Comprehensive Psychiatry*, 55(4), 1055-1062. <https://doi.org/10.1016/j.comppsy.2014.01.008>
- Linehan, K., Fennell, K. M., Hughes, D. L., & Wilson, C. J. (2017). Use of the Distress Thermometer in a cancer helpline context: Can it detect changes in distress, is it acceptable to nurses and callers, and do high scores lead to internal referrals? *European Journal of Oncology Nursing*, 26(Journal Article), 49-55. <https://doi.org/10.1016/j.ejon.2016.12.005>
- Liu, L., Wang, R., Sun, Y., Xiao, Y., Du, G., & Zhang, Q. (2024). Study on the cut-off point and the influencing factors of distress in newly diagnosed breast cancer patients. *Frontiers in Psychology*, 15(Journal Article). <https://doi.org/10.3389/fpsyg.2024.1281469>
- Ma, X., Zhang, J., Zhong, W., Shu, C., Wang, F., Wen, J., Zhou, M., Sang, Y., Jiang, Y., & 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. <https://doi.org/10.1007/s00520-014-2143-1>
- Martínez, P., Andreu, Y., Galdón, M. J., & Ibáñez, E. (2015). Improving the diagnostic accuracy of the Distress Thermometer: A potential role for the Impact Thermometer. *Journal of pain and symptom management*, 50(1), 124-129. <https://doi.org/10.1016/j.jpainsymman.2015.02.004>
- McElroy, J. A., Waindim, F., Weston, K., & Wilson, G. (2022). A systematic review of the translation and validation methods used for the national comprehensive cancer network distress thermometer in non-English speaking countries. *Psycho-Oncology*, 31(8), 1267-1274. <https://doi.org/10.1002/pon.5989>
- 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. <https://doi.org/10.1200/JCO.2006.10.0438>

- Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 27(5), 1171-1179. <https://doi.org/10.1007/s11136-017-1765-4>
- National Comprehensive Cancer Network. (2026, enero 14). *Distress Management* [Organization]. https://www.nccn.org/professionals/physician_gls/pdf/distress.pdf
- Nguyen T.Q., Do T.M., & Pham T.A. (2021). Screening for Psychological Distress in Vietnamese Cancer Patients: An Evaluation of the Distress Thermometer. *Cancer Medicine*, 10(21), 7793-7803. <https://doi.org/10.1002/cam4.4298>
- Olesen, M. L., Hansen, M. K., Hansson, H., Ottesen, B., Andersen, K. K., & Zoffmann, V. (2018). The distress thermometer in survivors of gynaecological cancer: Accuracy in screening and association with the need for person-centred support. *Supportive Care in Cancer*, 26(4), 1143-1150. <https://doi.org/10.1007/s00520-017-3935-x>
- OpenAI. (2026). *ChatGPT image generation* [Modelo de inteligencia artificial generativa]. <https://chatgpt.com/>
- Organización Mundial de la Salud. (2024, febrero 1). *Global cancer burden growing amidst mounting need for services* [Gubernamental]. <https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services>
- Ownby, K. K. (2019). Use of the Distress Thermometer in Clinical Practice. *Journal of the Advanced Practitioner in Oncology*, 10(2), 175-179.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Alonso-Fernández, S. (2021). Declaración PRISMA 2020: Una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología*, 74(9), 790-

799. <https://doi.org/10.1016/j.recesp.2021.06.016>

- Ploos van Amstel F.K., Tol J, Sessink K.H., van der Graaf W.T.A., Prins JB, & Ottevanger PB. (2017). A Specific Distress Cutoff Score Shortly After Breast Cancer Diagnosis. *Cancer Nursing*, 40(3), E35-E40. <https://doi.org/10.1097/NCC.0000000000000380>
- Priede, A., Ruiz Torres, M., Hoyuela, F., Herrán, A., & González-Blanch, C. (2014). El termómetro del estrés como prueba de cribado de ansiedad y depresión en pacientes oncológicos recién diagnosticados. *Psicooncología*, 11(1), 31-43. https://doi.org/https://doi.org/10.5209/rev_PSIC.2014.v11.n1.44915
- Ratnasekera, N., Perera, I., Kandapola Arachchige, P., & Jayasekara, P. (2023). Initiation of distress screening in oral cancer clinics in Sri Lanka. *Psycho-oncology*, 32(7), 1013-1021. <https://doi.org/10.1002/pon.6098>
- Recklitis, C. J., Blackmon, J. E., & Chang, G. (2016). Screening young adult cancer survivors for distress with the Distress Thermometer: Comparisons with a structured clinical diagnostic interview. *Cancer*, 122(2), 296-303. <https://doi.org/10.1002/cncr.29736>
- Robbeson C, Hugenholtz-Wamsteker W, Meeus M, Devoogdt N, Nijs J, & De Groef A. (2019). Screening of physical distress in breast cancer survivors: Concurrent validity of the Distress Thermometer and Problem List. *European Journal of Cancer Care*, 28(1), e12880. <https://doi.org/10.1111/ecc.12880>
- Sánchez-Román, S., Mejía-Castrejón, J., Montelongo-Niño, D., Barrientos-Casarrubias, V., Alfaro-Espeje, I., Quiroz-Friedman, P., Pacheco-Arroyo, D., Lara-Solares, A., Rivera-Cuevas, G., & Chiquete, E. (2020). Estrés en el cuidador primario del paciente adulto con enfermedad oncológica avanzada. *Latin American Journal of Clinical Sciences and Medical Technology*, 2, 1-8.
- Seib, C., Harbeck, E., Anderson, D., Porter-Steele, J., Nehill, C., Sanmugarajah, J., Perrin, L., Shannon, C., Cabraal, N., Jennings, B., Otton, G., Adams, C., Mellon, A., & Chambers, S. (2024). Establishing the sensitivity and specificity of the gynaecological cancer distress

- screen. *Psycho-oncology*, 33(3), 1-10. <https://doi.org/10.1002/pon.6328>
- Thalén-Lindström, A., Larsson, G., Hellbom, M., Glimelius, B., & Johansson, B. (2013). Validation of the Distress thermometer in a Swedish population of oncology patients; accuracy of changes during six months. *European Journal of Oncology Nursing*, 17(5), 625-631. <https://doi.org/10.1016/j.ejon.2012.12.005>
- Van Lander, A., Tarot, A., Savanovitch, C., Pereira, B., Vennat, B., & Guastella, V. (2019). Assessing the validity of the clinician-rated distress thermometer in palliative care. *BMC Palliative Care*, 18(1), 81. <https://doi.org/10.1186/s12904-019-0466-x>
- Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., Leeflang, M. M. G., Sterne, J. A. C., & Bossuyt, P. M. M. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529-536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3(1), 32-35. [https://doi.org/10.1002/1097-0142\(1950\)3:1%3C32::aid-cnrcr2820030106%3E3.0.co;2-3](https://doi.org/10.1002/1097-0142(1950)3:1%3C32::aid-cnrcr2820030106%3E3.0.co;2-3)
- Zheng, B., Du, P., Yi, T., Liu, J., Zeng, Z., Luo, D., & Jiang, Y. (2019). Effects of two translated phrases of distress thermometer on screening distress in Chinese cancer patients: A comparative study. *Journal of Clinical Nursing*, 28(5-6), 828-835. <https://doi.org/10.1111/jocn.14678>
- Zuber, C., Swartz, K., Chapman, A. E., Rhodes, L., Zhan, T. T., & Wen, K.-Y. (2025). Evaluation of the Distress Thermometer in older patients with cancer. *Journal of Geriatric Oncology*, 16(4). <https://doi.org/10.1016/j.jgo.2025.102221>

Anexo A

Figura A1

Termómetro del estrés, versión 2025



National
Comprehensive
Cancer
Network®

NCCN Guidelines, versión 1.2025

Manejo de la angustia

[Índice de NCCN Guidelines](#)
[Índice](#)
[Debate](#)

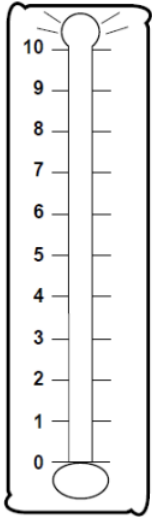
TERMÓMETRO DE ANGUSTIA DE NCCN

La angustia es una experiencia desagradable de naturaleza mental, física, social o espiritual. Puede afectar la forma en que piensa, siente o actúa. La angustia puede hacer que sea más difícil sobrellevar el cáncer, sus síntomas o su tratamiento.

Instrucciones: Marque con un círculo el número (0 a 10) que mejor describa la cantidad de angustia que experimentó la semana pasada, incluido el día de hoy.

Angustia extrema

Sin angustia



LISTA DE PROBLEMAS

En la última semana (incluido hoy), ¿ha tenido alguna inquietud relacionada con los elementos mencionados a continuación? (seleccione todos los elementos que apliquen).

Inquietudes sobre cuestiones físicas

- ☐ Dolor
- ☐ Dormir
- ☐ Fatiga
- ☐ Uso de tabaco
- ☐ Uso de sustancias ilícitas
- ☐ Memoria o concentración
- ☐ Salud sexual
- ☐ Cambios en la alimentación
- ☐ Pérdida o cambio de habilidades físicas

Inquietudes emocionales

- ☐ Preocupación o ansiedad
- ☐ Tristeza o depresión
- ☐ Pérdida de interés o disfrute
- ☐ Duelo o pérdida
- ☐ Miedo
- ☐ Soledad
- ☐ Enojo
- ☐ Cambios en la apariencia
- ☐ Sentimiento de inutilidad o de ser una carga

Inquietudes sociales

- ☐ Relación con su cónyuge o pareja
- ☐ Relación con sus hijos
- ☐ Relación con sus familiares
- ☐ Relación con sus amigos o compañeros de trabajo
- ☐ Comunicación con su equipo de atención médica
- ☐ Capacidad para tener hijos
- ☐ Prejuicio o discriminación

Inquietudes prácticas

- ☐ Cuidar de mí mismo
- ☐ Cuidar de otros
- ☐ Seguridad
- ☐ Trabajo
- ☐ Escuela
- ☐ Vivienda/servicios públicos
- ☐ Finanzas
- ☐ Seguro
- ☐ Transporte
- ☐ Cuidado infantil
- ☐ Tener suficiente alimento
- ☐ Acceso a medicamentos
- ☐ Decisiones de tratamiento

Inquietudes espirituales o religiosas

- ☐ Sentido o propósito
- ☐ Cambios de fe o creencias
- ☐ Muerte, morir o la vida después de la muerte
- ☐ Conflicto entre sus creencias y los tratamientos contra el cáncer
- ☐ Relación con lo sagrado
- ☐ Necesidades ritualistas o alimentarias

Otras inquietudes:

Nota: Todas las recomendaciones son de categoría 2A, salvo que se indique lo contrario.

Versión 1.2025, 18/11/2024 © 2024 National Comprehensive Cancer Network® (NCCN®). Todos los derechos reservados. Queda prohibida cualquier reproducción de NCCN Guidelines® y de esta ilustración sin el consentimiento expreso por escrito de NCCN.

DIS-A

Nota. Tomado de la National Comprehensive Cancer Network (2025)

Anexo B

Tabla B1

Componentes PIO y descriptores MeSH/DeCS

Componentes	P	I	O
MeSH	Neoplasms [Mesh]		"Reproducibility of Results"[Mesh] Validation Studies as Topic"[Mesh] Sensitivity and Specificity"[Mesh] Psychometrics"[Mesh]
DeCS	Neoplams		Psychometrics: Sensitivity and Specificity
Palabras clave			
En español	Cáncer Neoplasia	Termómetro del distrés	Validez, confiabilidad, sensibilidad, especificidad Propiedades psicométricas Precisión diagnóstica

Nota: Filtro: adultos; 10 últimos años

Anexo C

Tabla C1

Estrategia de búsqueda en cinco bases de datos

Base de datos	Estrategia de búsqueda	Resultados
PubMed	((("neoplasms"[MeSH Terms] OR "cancer"[Title/Abstract]) AND "distress thermometer"[Title/Abstract] AND ("reproducibility of results" [MeSH Terms] OR "validation studies as topic"[MeSH Terms] OR "sensitivity and specificity" [MeSH Terms] OR "psychometrics"[MeSH Terms] OR "data accuracy" [MeSH Terms] OR "validity" [Title/Abstract] OR "validation"[Title/Abstract] OR "reliability" [Title/Abstract] OR "sensitivity" [Title/Abstract] OR "specificity" [Title/Abstract] OR "psychometric properties" [Title/Abstract] OR "psychometric" [Title/Abstract] OR "diagnostic accuracy" [Title/Abstract])) AND ((y_10[Filter]) AND (alladult[Filter]))	92
SciELO	((cancer) OR (neoplasms)) AND (distress thermometer)	7
PsycINFO	((abstract: (neoplasm)) OR (abstract: (cancer))) AND ((abstract: ("distress thermometer"))) AND (((abstract: ("psychometric properties")) OR (abstract: (reliability)) OR (abstract: (validity)) OR (abstract: (sensitivity)) OR (abstract: ("diagnostic accuracy")))) AND ((Year: [2015 TO 2025]) AND ((AgeGroupFilt: ("Adulthood (18 yrs & older)")))))	37
LILACS	((validity) OR (reliability) OR (validation) OR (sensitivity) OR (specificity) OR (diagnostic accuracy) OR (psychometrics) OR (psychometric properties)) AND ((distress thermometer)) AND ((cancer) OR (neoplasms))) AND instance:"lilacsplus" AND db:"LILACS") AND instance:"lilacsplus"	5
COCHRANE LIBRARY	Search Name: Date Run: 07/08/2025 06:29:24 Comment: ID Search Hits #1 neoplasms 111089 #2 ("Cáncer"): ti, ab, kw OR (neoplasms): ti, ab, kw 243774 #3 #1 OR #2 243997 #4 (distress thermometer): ti, ab, kw 509 #5 MeSH descriptor: [Reproducibility of Results] explode all trees 16705. #6 (validity): ti, ab, kw OR ("validation study"): ti, ab, kw OR ("validation studies") ti, ab, kw OR ("reliability"): ti, ab, kw OR ("reliability coefficient") ti, ab, kw 27348 #7 ("sensitivity and specificity"): ti, ab, kw OR ("sensitivity"): ti, ab, kw OR ("specificity"): ti, ab, kw OR ("psychometric"): ti, ab, kw 85099 #8 #5 OR #6 OR #7 114315 #9 #3 AND #4 AND #8 20	20
		161

Anexo D

Figura D1

Países de procedencia de los estudios sobre el termómetro del estrés en adultos con cáncer



Nota: Elaboración propia. La figura fue elaborada por los autores con apoyo de herramientas de inteligencia artificial generativa únicamente para la visualización gráfica.

Anexo E

Tabla E1

Características de los estudios incluidos

Autor (año)	Diseño	n	Edad	País	Contexto clínico	Diagnóstico	Estadía	Tratamiento	Modalidad	Evaluador	Instrumento/ criterio de referencia
Alosaimi et al., 2018	Híbrido	247	18–77 años	Arabia Saudita	Ambulatorio	Mixto	I–IV	No reportado	Lápiz y papel	Médicos y residentes	HADS
Andreu et al., 2019	Híbrido	236	$M = 58$, $SD = 13.4$	España	Ambulatorio	Mixto	No reportado	Tratamiento en curso	Lápiz y papel	Enfermería	BSI-18
Antoniadis et al., 2015	Híbrido	84	18–82 años	Grecia	Hospitalizado	Cáncer de colon	I–IV	Prequirúrgico	Lápiz y papel	Autorreporte	HADS
Ascensio et al., 2024	Precisión diagnóstica	356	46–65 años	México	Cuidados paliativos	Mixto	Avanzado	Cuidados paliativos	No reportado	Psicólogos	ESAS-r
Bergerot et al., 2024	Psicométrico	302	$M = 55$, $SD = 14.4$	Brasil	Ambulatorio	Mixto	I–IV	Sin tratamiento	Lápiz y papel	Autorreporte	SupportScreen
Boudin et al., 2022	Precisión diagnóstica	93	70–95 años	Francia	Ambulatorio	Mixto	Mixto	Mixto	Lápiz y papel	Profesional de salud	DSM-5
Chiou et al., 2017	Precisión diagnóstica	768	18–92 años	Taiwán	Hospitalizado	Mixto	I–IV	Mixto	Lápiz y papel	Profesional de salud	CHQ-12
Civilotti et al., 2020	Precisión diagnóstica	436	22–87 años	Italia	Ambulatorio	Cáncer de mama	No reportado	Prequirúrgico	Lápiz y papel	Enfermería	HADS

Conev et al., 2019	Precisión diagnóstica	262	29–81 años	Bulgaria	Ambulatorio	Mixto	II–IV	Prequimioterapia	Lápiz y papel	Autorreporte	Medida repetida
Ghazali et al., 2017	Precisión diagnóstica	261	57–77 años	Reino Unido	Ambulatorio	Cáncer de cabeza y cuello	I–IV	Mixto	Lápiz y papel	Autorreporte	UWQOL
Guan et al., 2019	Precisión diagnóstica	441	18–80 años	China	Hospitalizado	Mixto	Avanzado	Tratamiento del dolor	Lápiz y papel	Profesional de salud	HADS
Kuchen et al., 2025	Precisión diagnóstica	763	$M = 53$, $SD = 13.7$	Alemania	Ambulatorio	Glioma	I y IV	No reportado	Lápiz y papel	Autorreporte	EF
Kyranou et al., 2020	Precisión diagnóstica	152	$M = 58.9$, $SD = 12.1$	Grecia	Ambulatorio	Mixto	No reportado	Quimioterapia	Lápiz y papel	Enfermería	MINI (CIE-10)
Linehan et al., 2017	Psicométrico	100	22–82 años	Australia	Línea telefónica	Mixto	No reportado	No reportado	Online	Enfermería	DASS-21
Liu et al., 2024	Precisión diagnóstica	167	$M = 46.9$, $SD = 8.59$	China	Hospitalizado	Cáncer de mama	No reportado	Prequirúrgico	Lápiz y papel	Autorreporte	HADS
Martinez et al., 2015	Precisión diagnóstica	385	$M = 58.7$, $SD = 13.2$	España	Ambulatorio	Mixto	I–IV	Mixto	Lápiz y papel	Autorreporte	BSI-18
Nguyen et al., 2021	Precisión diagnóstica	300	18–80 años	Vietnam	Ambulatorio	Mixto	Mixto	Mixto	Lápiz y papel	Autorreporte	PHQ-9 / GAD-7
Olesen et al., 2018	Precisión diagnóstica	165	38–61 años	Dinamarca	Online	Cáncer ginecológico	Temprano	Cirugía	Online	Autorreporte	HADS

Ploos Van Amstel et al., 2017	Precisión diagnóstica	181	30–87 años	Países Bajos	Ambulatorio	Cáncer de mama	No reportado	Mixto	Lápiz y papel	Autorreporte	HADS
Ratnasekera et al., 2023	Híbrido	95	60–89 años	Sri Lanka	Ambulatorio	Cáncer oromaxilofacial	I–IV	Tratamiento activo	Lápiz y papel	Psiquiatras	DSM-IV
Recklitis et al., 2016	Precisión diagnóstica	247	18–40 años	Estados Unidos	Ambulatorio	Supervivientes	No reportado	Mixto	Lápiz y papel	Autorreporte	DSM-IV (SCID)
Robbeson et al., 2019	Psicométrico	90	37–80 años	Países Bajos	Ambulatorio	Cáncer de mama	No reportado	Mixto	Lápiz y papel	Autorreporte	SF-36
Sarenmalin et al., 2018	Psicométrico	488	32–93 años	Suecia	Ambulatorio	Cáncer colorrectal	No metastásico	Prequirúrgico	Lápiz y papel	Autorreporte	QLQ-C30
Seib et al., 2024	Precisión diagnóstica	373	19–91 años	Australia	Hospitalario	Cáncer ginecológico	I–IV / recurrencia	Mixto	Lápiz y papel	Autorreporte	K-10
Van Lander et al., 2019	Híbrido	51	36–90 años	Francia	Cuidados paliativos	Mixto	Metastásico	Mixto	Lápiz y papel	Autorreporte	Criterio clínico
Zheng et al., 2019	Híbrido	172	$M = 53.5$, $SD = 12$	China	Hospitalizado	Tumores sólidos	No reportado	Quimioterapia	Lápiz y papel	Autorreporte	HADS
Zuber et al., 2025	Precisión diagnóstica	242	65–90 años	Estados Unidos	Ambulatorio	Mixto	No reportado	Quimioterapia	Lápiz y papel	Autorreporte	GDS

Nota: Elaboración propia

Anexo F

Tabla F1

Evaluación de riesgo de sesgo según QUADAS-2

Autor (año)	D1 Riesgo	D1 Aplic.	D2 Riesgo	D2 Aplic.	D3 Riesgo	D3 Aplic.	D4 Riesgo	D4 Aplic.
Alosaimi et al., 2018	●	●	●	●	●	●	●	●
Andreu et al., 2019	●	●	●	●	●	●	●	●
Antoniadis et al., 2015	●	●	●	●	●	●	●	●
Ascencio et al., 2024	●	●	●	●	●	●	●	●
Boudin et al., 2022	●	●	●	●	●	●	●	●
Chiou et al., 2017	●	●	●	●	●	●	●	●
Civilotti et al., 2020	●	●	●	●	●	●	●	●
Ghazali et al., 2017	●	●	●	●	●	●	●	●
Guan et al., 2019	●	●	●	●	●	●	●	●
Kuchen et al., 2025	●	●	●	●	●	●	●	●
Kyranou et al., 2020	●	●	●	●	●	●	●	●
Liu et al., 2024	●	●	●	●	●	●	●	●
Martinez et al., 2015	●	●	●	●	●	●	●	●
Nguyen et al., 2021	●	●	●	●	●	●	●	●
Olesen et al., 2018	●	●	●	●	●	●	●	●
Ploos Van Amstel et al., 2017	●	●	●	●	●	●	●	●
Ratnasekera et al., 2023	●	●	●	●	●	●	●	●
Recklitis et al., 2016	●	●	●	●	●	●	●	●
Robbeson et al., 2019	●	●	●	●	●	●	●	●
Seib et al., 2024	●	●	●	●	●	●	●	●
Van Lander et al., 2019	●	●	●	●	●	●	●	●
Zheng et al., 2019	●	●	●	●	●	●	●	●
Zuber et al., 2025	●	●	●	●	●	●	●	●

Nota: D1 = selección de pacientes; D2 = prueba índice; D3 = estándar de referencia; D4 = flujo y tiempo; Aplic. = aplicabilidad. ● = bajo riesgo o baja preocupación; ● = riesgo o preocupación inciertos; ● = alto riesgo o alta preocupación.

Anexo G

Tabla G1

Evaluación de riesgo de sesgo según COSMIN

Autor (año)	Propiedad	Box COSMIN	Calidad metodológica	Resultado	Juicio	Justificación
Alosaimi et al., 2018	Validez de constructo (convergente)	9	Adecuada	$r = .50$; AUC = .76	+	Correlación moderada con HADS y capacidad discriminativa aceptable.
Andreu et al., 2019	Validez de constructo (convergente)	9	Adecuada	AUC = .70	+	Buen desempeño global frente a BSI-18; punto de corte ≥ 5 útil para cribado.
Antoniadis et al., 2015	Validez de constructo (convergente)	9	Adecuada	$r = .39$; AUC = .81	+	Correlación significativa con HADS y adecuada discriminación.
Bergerot et al., 2024	Validez de constructo (convergente)	9	Muy adecuada	$r = .80$	+	Alta convergencia con dimensiones emocionales del SupportScreen.
Conev et al., 2019	Validez de constructo (hipotética/discriminativa)	9	Dudosa	$r = -.19$; AUC = .60	-	Asociación débil y baja capacidad discriminativa.
Linehan et al., 2017	Validez de constructo (convergente)	9	Adecuada	$r = .45-.64$	+	Correlaciones consistentes con depresión, ansiedad y estrés (DASS-21).
Nguyen et al., 2021	Validez de constructo (convergente)	9	Adecuada	$r = .53$; AUC = .81-.82	+	Correlaciones moderadas y buen rendimiento discriminativo.
Olesen et al., 2018	Validez de constructo (ROC-based)	9	Adecuada	AUC = .73	+	Discriminación aceptable; alta sensibilidad para cribado.
Ratnasekera et al., 2023	Fiabilidad (acuerdo)	6	Adecuada	$\kappa = .89$	+	Acuerdo casi perfecto en clasificación diagnóstica.
Robbeson et al., 2019	Validez de constructo (convergente inversa)	9	Adecuada	$r = -.45$	+	Relación inversa coherente con calidad de vida.
Sarenmalm et al., 2018	Validez de constructo (convergente)	9	Adecuada	$r = .74$	+	Fuerte asociación con dominio emocional del QLQ-C30.
Zheng et al., 2019	Validez de constructo (convergente)	9	Adecuada	$r = .78$	+	Alta correlación con HADS.
Van Lander et al., 2019	Fiabilidad (acuerdo inter-evaluador)	6	Adecuada	$\kappa = .52$	?	Concordancia moderada entre clínico y paciente; variabilidad entre evaluadores.
Zuber et al., 2025	Validez de constructo (convergente)	9	Adecuada	$r = .41$; AUC = .70	+	Correlación moderada con GDS y discriminación limitada.

Nota: + = resultado favorable; - = resultado desfavorable = resultado indeterminado o evidencia insuficiente para establecer un juicio concluyente.

Anexo H

Tabla H1

Capacidad de discriminación del TD aplicando el índice de Youden

Autor (año)	Sensibilidad	Especificidad	Punto de corte	Índice de Youden (J)
Alosaimi et al., 2018	.70	.63	≥4	.33
Andreu et al., 2019	.89	.59	≥5	.48
Antoniadis et al., 2015	.73	.79	≥5	.52
Ascencio et al., 2024	.76	.54	≥5	.30
Boudin et al., 2022	.50	.97	≥4	.47
Chiou et al., 2017	.72	.80	≥4	.52
Civilotti et al., 2020	.89	.64	≥4	.53
Ghazali et al., 2017	.84	.75	≥4	.59
Guan et al., 2019	.86	.53	≥5	.39
Kuchen et al., 2025	.81	.46	≥5	.27
Kyranou et al., 2020	.85	.69	≥4	.54
Liu et al., 2024	1.00	.72	≥4	.72
Martinez et al., 2015	.93	.57	≥5	.50
Nguyen et al., 2021 (PHQ-9)	.87	.58	≥4	.45
Nguyen et al., 2021 (GAD-7)	.92	.58	≥4	.50
Olesen et al., 2018	.93	.36	≥2	.29
Ploos Van Amstel et al., 2017	.73	.84	≥7	.57
Ratnasekera et al., 2023	.98	.78	≥4	.76
Recklitis et al., 2016	.68	.78	≥5	.46
Robbeson et al., 2019	.53	.68	≥4.5	.21
Seib et al., 2024	.92	.73	≥5	.65
Van Lander et al., 2019	.83	.71	≥3	.54
Zheng et al., 2019	.78	.79	≥3	.57
Zuber et al., 2025	.66	.68	≥4	.34

Nota: Elaboración propia