

# Herramientas para la detección precoz del malestar emocional en personas adultas con cáncer

Early detection tools  
for emotional distress in  
adult people with cancer

Informes de Evaluación de Tecnologías Sanitarias  
ACIS, Avalia-t

INFORMES, ESTUDIOS E INVESTIGACIÓN



# Herramientas para la detección precoz del malestar emocional en personas adultas con cáncer

Early detection tools  
for emotional distress in  
adult people with cancer

Informes de Evaluación de Tecnologías Sanitarias  
ACIS, Avalia-t

INFORMES, ESTUDIOS E INVESTIGACIÓN



AXENCIA GALEGA  
DE COÑECIMENTO  
EN SAÚDE

GÓMEZ SALGADO, PATRICIA

Herramientas para la detección precoz del malestar emocional en personas adultas con cáncer / Patricia Gómez Salgado ... [et al]. — Madrid: Ministerio de Sanidad; Santiago de Compostela: Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS, Unidad de Asesoramiento Científico-Técnico, Avalia-t; 2023.

1 archivo PDF — (Informes, Estudios e Investigación)

NIPO: 133-23-026-5

Depósito Legal: C 1214-2023

1. Distrés psicológico 2. Neoplasias 3. Sensibilidad y especificidad I. Triñanes Pego, Yolanda. II. Casal Acción, Beatriz. III. Faraldo Vallés, María José. IV. España. Ministerio de Sanidad. V. Agencia Gallega para la Gestión del Conocimiento, ACIS.

El contenido del presente informe es responsabilidad exclusiva de la Unidad de Asesoramiento Científico-técnico, Avalia-t, sin que la colaboración de las y los revisores externos presuponga por su parte la completa aceptación de este.

Este documento puede ser reproducido total o parcialmente, por cualquier medio, siempre que se cite explícitamente su procedencia.

Información dirigida a profesionales de la salud.

Edición: 2023

Edita: Ministerio de Sanidad.

Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS  
Unidad de Asesoramiento Científico-Técnico, Avalia-t.

NIPO: 133-23-026-5

Depósito legal: C 1214-2023

Contacto: Avalia\_t.Acis@sergas.es

Maquetación: Tórculo Comunicación Gráfica, S. A.

Este documento ha sido realizado por la Unidad de Asesoramiento Científico técnico, Avalia-t, de la Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS, en el marco de la financiación del Ministerio de Sanidad para el desarrollo de las actividades del Plan anual de Trabajo de la Red Española de Agencias de Evaluación de Tecnologías Sanitarias y Prestaciones del SNS, aprobado en el Pleno del Consejo Interterritorial del SNS del 26 de mayo de 2021.

Para citar este informe:

Gómez Salgado P, Triñanes Pegó Y, Casal Acción B, Faraldo Vallés MJ. Herramientas para la detección precoz del malestar emocional en personas adultas con cáncer. Madrid: Ministerio de Sanidad. Santiago de Compostela: Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS, Unidad de Asesoramiento Científico-técnico; Avalia-t; 2023.



# Índice

<b>Lista de abreviaturas</b> .....	7
<b>Lista de tablas</b> .....	10
<b>Lista de figuras</b> .....	11
<b>Resumen</b> .....	12
<b>Summary</b> .....	18
<b>Justificación</b> .....	23
<b>1. Introducción</b> .....	24
1.1. Descripción del problema de salud. ....	24
1.1.1. Problema de salud diana .....	24
1.1.2. Manejo actual del problema de salud .....	27
1.1.3. Población diana .....	29
1.2. Descripción y características técnicas de la tecnología a estudio .....	30
1.2.1. Características de la tecnología .....	30
1.2.2. Regulación: licencias y autorizaciones. ....	39
1.2.3. Utilización .....	40
1.2.4. Requerimientos de la técnica .....	41
<b>2. Alcance y objetivo</b> .....	43
2.1. Objetivo .....	43
2.2. Alcance .....	43
<b>3. Método</b> .....	44
3.1. Metodología de elaboración del informe .....	44
3.1.1. Estrategia de búsqueda bibliográfica .....	44
3.1.2. Criterios de selección de estudios. ....	45
3.1.3. Extracción de datos y síntesis de la evidencia. ....	46
3.1.4. Valoración de la calidad de la evidencia .....	47
3.1.5. Metodología para la incorporación de la visión de los y las pacientes .....	48
<b>4. Resultados</b> .....	49
4.1. Descripción de la evidencia disponible .....	49
4.1.1. Resultados de la búsqueda .....	49
4.1.2. Descripción de las guías de práctica clínica seleccionadas .....	50
4.1.3. Descripción y calidad de los estudios seleccionados .....	51
4.2. Efectividad clínica de la tecnología .....	62
4.2.1. Propiedades psicométricas de las herramientas .....	62

4.2.2. Precisión diagnóstica de las herramientas. . . . .	68
4.2.3. Valoración global de las escalas . . . . .	75
4.2.4. Síntesis global de la evidencia. . . . .	75
<b>5. Consideraciones de implementación . . . . .</b>	<b>77</b>
5.1. Aspectos organizativos . . . . .	77
5.2. Aspectos éticos . . . . .	78
5.3. Aspectos sociales y relacionados con la visión de los y las pacientes. . . . .	79
<b>6. Discusión. . . . .</b>	<b>81</b>
6.1. Discusión de la metodología. . . . .	81
6.2. Discusión de los resultados de efectividad . . . . .	82
6.3. Discusión de los aspectos organizativos, éticos, sociales y relacionados con la visión de los y las pacientes . . . . .	85
<b>7. Conclusiones . . . . .</b>	<b>88</b>
<b>Contribuciones . . . . .</b>	<b>89</b>
Contribución de las autoras. . . . .	89
Revisión externa . . . . .	89
Secretaría . . . . .	90
Agradecimientos . . . . .	90
Declaración de intereses . . . . .	90
<b>Referencias bibliográficas . . . . .</b>	<b>91</b>
<b>Anexos . . . . .</b>	<b>103</b>
Anexo A. Estrategia de búsqueda . . . . .	103
Anexo B. Tablas de evidencia . . . . .	110
Anexo C. Referencias excluidas. . . . .	136
Anexo D. Evaluación de la calidad de los estudios de precisión diagnóstica: herramienta QUADAS-2. . . . .	139
Anexo E. Valoración individual del riesgo de sesgo y la aplicabilidad de cada estudio mediante la herramienta QUADAS-2. . . . .	141

# Lista de abreviaturas

<b><math>\alpha</math>:</b>	Alfa de Cronbach
<b>ACIS:</b>	Agencia Gallega para la Gestión del Conocimiento en Salud
<b>AECC:</b>	Asociación Española Contra el Cáncer
<b>AFC:</b>	Análisis factorial confirmatorio
<b>AFE:</b>	Análisis factorial exploratorio
<b>AIC:</b>	Akaike Information Criterion
<b>AMSTAR-2:</b>	A Measurement Tool to Assess Systematic Reviews
<b>An:</b>	Ansiedad
<b>APA:</b>	Asociación Estadounidense de Psicología
<b>ASCO:</b>	Sociedad Estadounidense de Oncología Clínica
<b>AUC:</b>	Área bajo la curva
<b>BSI-18:</b>	Brief Symptom Inventory 18 Inventario Breve de Síntomas 18
<b>CanCon:</b>	Acción Conjunta Europea para el Control del Cáncer
<b>CEXT:</b>	Consulta externa
<b>CFI:</b>	Comparative Fit Index
<b>CINAHL:</b>	Cumulative Index to Nursing & Allied Health
<b>COP:</b>	Consejo General de la Psicología de España
<b>CR:</b>	Fiabilidad compuesta
<b>CUI-:</b>	Utilidad clínica negativa
<b>CUI+:</b>	Utilidad clínica positiva
<b>De:</b>	Depresión
<b>DEG:</b>	Distrés Emocional Global
<b>DME:</b>	Cuestionario de Detección de Malestar Emocional
<b>DSM-IV:</b>	Manual diagnóstico y estadístico de los trastornos mentales (4ª ed.)
<b>DT:</b>	Distress Thermometer Termómetro de distrés o malestar
<b>dt:</b>	Desviación típica
<b>EBSCO:</b>	Elton B. Stephens Company
<b>EE.UU.:</b>	Estados Unidos

<b>ENP-E:</b>	Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida
<b>EORTC:</b>	Organización Europea para la Investigación y el Tratamiento del Cáncer
<b>ESAS:</b>	Edmonton Symptom Assessment System Cuestionario de evaluación de síntomas de Edmonton
<b>ESAS-r:</b>	Edmonton Symptom Assessment System Scale Revised Cuestionario de evaluación de síntomas de Edmonton, versión revisada
<b>ESMO:</b>	Sociedad Europea de Oncología Médica
<b>ETS:</b>	Evaluación de tecnologías sanitarias
<b>EVA:</b>	Escalas Visuales Analógicas
<b>FIVO:</b>	Fundación Instituto Valenciano de Oncología
<b>FLC:</b>	Fichas de Lectura Crítica
<b>GAD-7:</b>	Generalized Anxiety Disorder-7 item Escala de Ansiedad Generalizada (versión de 7 ítems)
<b>G-I-N:</b>	Red Internacional de Guías
<b>gl:</b>	Grados de libertad
<b>GPC:</b>	Guías de práctica clínica
<b>GSI:</b>	Índice Global de Severidad
<b>HADS:</b>	Hospital Anxiety and Depression Scale Escala Hospitalaria de Ansiedad y Depresión
<b>IC:</b>	Intervalo de confianza
<b>ICC:</b>	Índice de correlación intraclase
<b>ICO:</b>	Instituto Catalán de Oncología
<b>IFI:</b>	Incremental Fit Index
<b>IT:</b>	Impact Thermometer Termómetro de impacto
<b>KMO:</b>	Índice Kaiser-Mayer-Olkin
<b>MMSE:</b>	Mini-Mental State Examination Mini-Examen del Estado Mental
<b>MSPSS:</b>	Multidimensional Scale of Perceived Social Support Escala Multidimensional de Apoyo Social Percibido
<b>NCCN®:</b>	Red Nacional Integral sobre el Cáncer de EE.UU.
<b>NICE:</b>	Instituto Nacional para la Salud y la Excelencia Clínica del Reino Unido

<b>NNFI:</b>	Non-Normed Fit Index
<b>OMS:</b>	Organización Mundial de la Salud
<b>PIRDS:</b>	Población objetivo, prueba índice, prueba de referencia, condición de interés y diseño de estudios
<b>PHQ-2:</b>	Patient Health Questionnaire-2 Cuestionario de Salud del Paciente-2
<b>PHQ-9:</b>	Patient Health Questionnaire-9 Cuestionario de Salud del Paciente-9
<b>PL:</b>	Problem List Lista de Problemas
<b>PPS:</b>	Palliative Performance Scale Escala funcional paliativa
<b>QLQ-C15-PAL:</b>	Quality of Life Questionnaire Core 15 Palliative Cuestionario de Calidad de Vida Core 15 Palliative
<b>QUADAS-2:</b>	Tool for the Quality Assessment of Diagnostic Accuracy Studies
<b>RCFI:</b>	Robust Comparative Fit Index
<b>RedETS:</b>	Red Española de Agencias de Evaluación de Tecnologías Sanitarias y Prestaciones del Sistema Nacional de Salud
<b>RMSEA:</b>	Root Mean Square Error of Approximation
<b>ROC:</b>	Receiver Operating Characteristic
<b>RS:</b>	Revisión sistemática
<b>S-B<math>\chi^2</math>:</b>	Chi cuadrado de Satorra Bentler
<b>SCID-I:</b>	Entrevista clínica estructurada para los trastornos del eje I del DSM-IV
<b>SCP:</b>	Servicio de Cuidados Paliativos
<b>SDS:</b>	Self-rating Depression Scale Escala de depresión de Zung
<b>SEOM:</b>	Sociedad Española de Oncología Médica
<b>SNS:</b>	Sistema Nacional de Salud
<b>SRMR:</b>	Standardized Root Mean Square Error of the Residuals
<b>STAI-S:</b>	Inventario de Ansiedad Estado-Rasgo (subescala de ansiedad-estado)
<b>TLI:</b>	Tucker-Lewis Index
<b>VPN:</b>	Valor predictivo negativo
<b>VPP:</b>	Valor predictivo positivo
<b>WoS:</b>	Web of Science

# Lista de tablas

Tabla 1.	Continuo del malestar emocional . . . . .	24
Tabla 2.	Factores de riesgo de malestar emocional . . . . .	25
Tabla 3.	Ventajas y desventajas de las herramientas de cribado, según su longitud. . . . .	31
Tabla 4.	Cuadro resumen de las GPC analizadas y las herramientas de detección precoz del malestar emocional en entorno clínico recomendadas por cada una de ellas . . .	32
Tabla 5.	Ventajas e inconvenientes del cribado con lápiz y papel versus cribado digital . . . . .	42
Tabla 6.	Pregunta PIRDS (población objetivo, prueba índice, prueba de referencia, condición de interés y diseño de estudios) . . . . .	43
Tabla 7.	Bases de datos bibliográficas de literatura biomédica consultadas. . . . .	44
Tabla 8.	Características de las herramientas de detección precoz analizadas . . . . .	53
Tabla 9.	Descripción contextual de los estudios de validación seleccionados . . . . .	56
Tabla 10.	Consistencia interna y evidencias de validez del BSI-18 en los estudios españoles de validación en población oncológica . . . . .	63
Tabla 11.	Consistencia interna y evidencias de validez de la HADS en los estudios españoles de validación en población oncológica . . . . .	65
Tabla 12.	Evidencias de validez del DT en los estudios españoles de validación en población oncológica . . . . .	66
Tabla 13.	Consistencia interna y evidencias de validez del ENP-E en los estudios españoles de validación en población oncológica . . . . .	67
Tabla 14.	Precisión diagnóstica del BSI-18, según los estudios de validación española en población oncológica . . . . .	69
Tabla 15.	Precisión diagnóstica del ESAS-r, según el estudio de validación española en población oncológica . . . . .	69
Tabla 16.	Precisión diagnóstica de las EVA de ansiedad y depresión, según el estudio de validación española en población oncológica . . . . .	70
Tabla 17.	Precisión diagnóstica de la HADS, según los estudios de validación española en población oncológica . . . . .	71
Tabla 18.	Precisión diagnóstica del DT, según los estudios de validación española en población oncológica . . . . .	72
Tabla 19.	Precisión diagnóstica del DME, según el estudio de validación española en población oncológica . . . . .	73
Tabla 20.	Precisión diagnóstica del ENP-E, según el estudio de validación española en población oncológica . . . . .	74
Tabla 21.	Valoración global de las herramientas, según el procedimiento de Vodermaier <i>et al.</i> (48) . . . . .	75
Tabla 22.	Resumen de las características más relevantes de los estudios de validación y las herramientas de detección precoz analizados. . . . .	76

# Lista de figuras

Figura 1.	Versión española del Distress Thermometer . . . . .	33
Figura 2.	Cuestionario Edmonton Symptom Assessment System-Revised (ESAS-r) . . . . .	34
Figura 3.	Versión española del Patient Health Questionnaire-2 (PHQ-2). . . . .	35
Figura 4.	Canadian Problem Checklist . . . . .	36
Figura 5.	Versión española del Patient Health Questionnaire-9 (PHQ-9). . . . .	37
Figura 6.	Versión española de la Generalized Anxiety Disorder-7 (GAD-7) . . . . .	38
Figura 7.	Diagrama de flujo de los resultados de la búsqueda inicial de guías de práctica clínica. . . . .	49
Figura 8.	Diagrama de flujo de los resultados de la búsqueda bibliográfica de estudios primarios . . . . .	50
Figura 9.	Evaluación de la calidad de los tres estudios sobre el BSI-18, según la herramienta QUADAS-2 . . . . .	58
Figura 10.	Evaluación de la calidad del estudio sobre el ESAS-r, según la herramienta QUADAS-2 . . . . .	59
Figura 11.	Evaluación de la calidad del estudio sobre las EVA de ansiedad y depresión, según la herramienta QUADAS-2 . . . . .	59
Figura 12.	Evaluación de la calidad de los tres estudios sobre la HADS, según la herramienta QUADAS-2 . . . . .	60
Figura 13.	Evaluación de la calidad de los cinco estudios sobre el DT, según la herramienta QUADAS-2 . . . . .	61
Figura 14.	Evaluación de la calidad del estudio sobre el DME, según la herramienta QUADAS-2 . . . . .	61
Figura 15.	Evaluación de la calidad del estudio sobre el ENP-E, según la herramienta QUADAS-2 . . . . .	62

# Resumen

**Introducción:** El modelo de atención a las personas con cáncer se está reformulando hacia un enfoque integral de la salud. Esto conlleva tener en cuenta que sus necesidades van más allá de los aspectos físicos relacionados con la enfermedad y su tratamiento, e incluyen un amplio abanico de implicaciones emocionales, interpersonales o sociales.

En este marco, la Red Nacional Integral sobre el Cáncer de EE.UU. (NCCN®) define el malestar o distrés emocional como una experiencia multifactorial desagradable de naturaleza psicológica, social, espiritual o física que puede interferir en la capacidad de afrontar de forma efectiva el cáncer, sus síntomas físicos y su tratamiento. Este malestar se extiende a lo largo de un continuo que comprende desde sentimientos de vulnerabilidad, tristeza y temor hasta problemas que pueden llegar a ser incapacitantes como la depresión, la ansiedad, el pánico, el aislamiento social y las crisis existenciales o espirituales.

Estudios recientes estiman que la prevalencia de distrés en población oncológica adulta oscila entre un 25 % y un 52 %. Conscientes de esta realidad prevalente, la Estrategia en Cáncer del Sistema Nacional de Salud (2021), dentro de la línea estratégica sobre el modelo asistencial, plantea como uno de sus objetivos utilizar en la atención oncológica una herramienta estandarizada para la detección precoz del malestar emocional (*distress*) validada al español.

**Objetivo:** Proporcionar una evaluación de las características y las propiedades psicométricas de las herramientas para la detección precoz del malestar emocional en población oncológica adulta validadas en el contexto español que puedan utilizarse en el contexto del Sistema Nacional de Salud (SNS).

**Métodos:** Tras una búsqueda bibliográfica preliminar de guías de práctica clínica (GPC) e informes de evaluación de tecnologías sanitarias (realizada en octubre de 2021), se diseñaron dos estrategias de búsqueda específicas en las principales bases de datos biomédicas: la primera, para localizar guías de práctica clínica que aborden el malestar emocional en pacientes oncológicos (noviembre 2021) y, la segunda, para localizar los estudios primarios de elaboración y validación de herramientas en el contexto español, así como aquellos de adaptación transcultural y validación de herramientas al contexto español (diciembre de 2021). Todo este proceso se completó mediante la creación de alertas en las principales bases de datos (Medline, Embase y WoS) con el fin de recuperar todos los estudios relevantes que pudiesen publicarse durante la redacción del informe. Las alertas se mantuvieron activas hasta la fecha de edición del informe. Tras la lectura de los títulos y los resúmenes de los artículos resultantes de la búsqueda, se realizó una selección de estudios conforme a los criterios de inclusión y exclusión previamente establecidos. Tanto la selección de estudios como la extracción de datos y el proceso de síntesis de la evidencia se ejecutó por pares. La información se sintetizó en tablas de evidencia elaboradas con la herramienta FLC 3.0 y adaptadas a los objetivos específicos de este informe. La calidad de la evidencia de los estudios se valoró mediante la herramienta QUADAS-2 para estudios de eficacia diagnóstica. Asimismo, se llevó a cabo una valoración global de cada una de las

escalas teniendo en cuenta la fiabilidad, el tipo de medida criterio empleada, el tamaño de las muestras y la validez de cada escala. La valoración sobre los sesgos y la aplicabilidad de los estudios, así como la valoración global de las escalas, se efectuó por pares. El método de resolución de desacuerdos establecido en todas las fases fue el consenso.

Para analizar los aspectos relacionados con la visión de los y las pacientes, se analizó la información correspondiente tanto en las GPC como en los estudios primarios incluidos. Además, este informe fue revisado por 3 representantes de pacientes de la Asociación Española Contra el Cáncer (AECC).

**Resultados:** Como resultado de la estrategia de búsqueda de guías, se tuvieron en cuenta 6 GPC para abordar los dominios relativos a la descripción del problema de salud, la descripción y las características técnicas de la tecnología, así como para analizar las consideraciones de implementación (aspectos organizativos, éticos y sociales, y relacionados con la visión de los y las pacientes). Como resultado de la búsqueda de estudios, se incluyeron 15 validaciones de herramientas de detección precoz del malestar emocional en pacientes en edad adulta con cáncer. Se evaluaron 7 escalas: las Escalas Visuales Analógicas (EVA) de ansiedad y depresión, el Termómetro de distrés (DT), el Cuestionario de Detección de Malestar Emocional (DME), el Inventario Breve de Síntomas<sup>18</sup> (BSI-18), el Cuestionario de evaluación de síntomas de Edmonton, versión revisada (ESAS-r), la Escala Hospitalaria de Ansiedad y Depresión (HADS) y el Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida (ENP-E). Según la longitud de la escala, 3 de las herramientas evaluadas (EVA de ansiedad y depresión, DT y DME) son ultracortas (concretamente, entre 1 o 2 ítems), y 4 de ellas (BSI-18, ESAS-r, HADS y ENP-E), cortas (específicamente, entre 10 y 18 ítems). Con respecto a la muestra de validación, 3 escalas (DT, BSI-18 y HADS) fueron validadas por tres o más estudios, contando con muestras heterogéneas de pacientes oncológicos en diferentes fases de la enfermedad; mientras que 4 escalas (EVA de ansiedad y depresión, DME, ESAS-r y ENP-E) fueron validadas a través de un único estudio para cada escala con pacientes oncológico-paliativos.

Según la herramienta QUADAS-2 para la evaluación de la calidad de los estudios de precisión diagnóstica, el estudio de validación de las EVA de ansiedad y depresión, cuatro de los cinco estudios de validación del DT, dos de los tres estudios de validación del BSI-18 y uno de los tres estudios de validación de la HADS no generan preocupaciones sobre sus riesgos de sesgo o sobre la aplicabilidad de sus resultados. El resto de los estudios genera ciertas preocupaciones, ya sea respecto al riesgo de sesgo o la aplicabilidad de sus resultados.

En lo que concierne a las propiedades psicométricas analizadas, los estudios de validación del DT no calculan ni su consistencia interna ni su estructura factorial, dado que está formado por un solo ítem. Por su parte, la validez convergente del test se evalúa en todos los estudios de validación, informando de correlaciones positivas moderadas con otras medidas como el BSI-18, los problemas emocionales recogidos en la Lista de Problemas, o la HADS. El estudio de validación de las EVA de ansiedad y depresión no informa ni de su consistencia interna ni de su estructura factorial, entendiéndolas como medidas

complementarias pero no pertenecientes a una misma escala. No obstante, sí analiza su validez convergente a través del cálculo de correlaciones positivas del Distrés Emocional Global (DEG) o la puntuación obtenida de la suma de las puntuaciones de las EVA con la puntuación total de la HADS, así como entre la EVA de ansiedad y la subescala de la HADS de ansiedad, y entre la EVA de depresión y la subescala de la HADS de depresión. Con respecto al DME, su consistencia interna ( $\alpha = 0,685$ ) resulta baja, aunque dicho valor se aproxima al mínimo aceptable ( $\alpha = 0,7$ ), sobre todo teniendo en cuenta que el instrumento está formado solo por dos ítems. El estudio de validación de esta herramienta informa de su validez convergente a través del cálculo de coeficientes de correlación intraclase (con el DT, la HADS y sus subescalas, así como con una valoración psicológica clínica), que resultan positivos. Por su parte, los estudios de validación del BSI-18 informan de una consistencia interna aceptable ( $\alpha$  entre 0,82 y 0,91), aportan evidencias de validez convergente y confirman una estructura jerárquica con tres factores de primer orden (subescalas de ansiedad, depresión y somatización) y un factor de segundo orden como medida de distrés emocional (el Índice Global de Severidad -GSI-). El ESAS-r también cuenta con una consistencia interna aceptable ( $\alpha = 0,86$ ), pero no analiza ni la validez de constructo ni la validez convergente de la escala. Con referencia a la HADS, los valores aportados por los estudios de validación sobre su consistencia interna son aceptables ( $\alpha > 0,80$ ) y la estructura subyacente encontrada en dos de los estudios es bifactorial, formada por las subescalas de ansiedad y depresión. El otro estudio de validación encuentra una solución jerárquica con dos factores de primer orden (ansiedad y depresión), y un factor de segundo orden (puntuación total) como medida de distrés emocional. Asimismo, uno de los estudios también aporta evidencias de validez convergente de cada subescala. En el caso del ENP-E, el estudio de validación informa de una consistencia interna aceptable ( $\alpha = 0,75$ ) y de una estructura cuatrifactorial subyacente, formada por las subescalas de bienestar emocional, apoyo social, espiritualidad e información. Con respecto a la validez convergente de la escala, resultan positivos los coeficientes de correlación calculados con la HADS y el DT. La capacidad discriminativa de cualquiera de las herramientas de detección precoz analizadas resultó ser buena (tomando el área bajo la curva valores superiores a 0,80 en todos los casos). En el caso del BSI-18 y el ESAS-r, no se calculó la capacidad discriminativa por falta de prueba de referencia.

La valoración global de cada escala se basó en una evaluación combinada de su fiabilidad, el tipo de medida criterio empleada, el tamaño de las muestras y la validez de cada escala (valor medio de su sensibilidad y especificidad). Siguiendo esta valoración multicriterio, tres escalas fueron valoradas como pobres (EVA de ansiedad y depresión, BSI-18 y ESAS-r), una escala fue valorada como aceptable (ENP-E) y tres escalas obtuvieron una valoración moderada (DT, DME y HADS).

El impacto organizativo de la introducción de herramientas de detección precoz del malestar emocional podría ser sustancial, especialmente si se introduce como parte de un programa de cribado o evaluación sistemática. Distintas GPC recomiendan la creación de equipos de trabajo que coordinen y lideren su implementación, teniendo entre sus miembros la figura de algún psicólogo o alguna psicóloga. Desde el punto de vista ético, el uso de una herramienta de detección precoz del malestar emocional cumple el

requerimiento de que el beneficio supere los posibles efectos secundarios. También se han de tener en cuenta otras consideraciones éticas como el respeto a la autonomía, el derecho a la intimidad y la confidencialidad y la equidad en el acceso.

En cuanto a la visión de los y las pacientes, apenas se ha identificado información sobre la aceptabilidad, la satisfacción o las perspectivas de las y los pacientes en las GPC recuperadas ni en los estudios analizados.

**Discusión:** En términos generales, la capacidad discriminativa de cada una de las herramientas analizadas es buena, aunque cabría matizar que los resultados coinciden con lo planteado por estudios previos, que apuntan que las medidas ultracortas suelen poseer una menor precisión diagnóstica, pues, aun teniendo adecuados niveles de sensibilidad, los valores de especificidad suelen ser de bajos a moderados. No obstante, estas medidas representan una ventaja en el ámbito sanitario, ya que son más sencillas de implementar dentro de la atención rutinaria.

A nivel general, el empleo de una u otra herramienta podrá diferir según el contexto y el tipo de pacientes con quienes vayan a ser empleadas. Para la elección, también ha de contemplarse la calidad de la evidencia de los estudios de validación de cada uno de los instrumentos, el número de estudios de validación que respalda cada herramienta y la valoración global de cada escala.

Teniendo en cuenta los resultados en su conjunto, encontramos 3 medidas ultracortas: dos para población paliativa (EVA de ansiedad y depresión, y cuestionario DME), y una para población oncológica general (Termómetro de distrés). En el caso de las EVA, la valoración de la calidad del único estudio de validación no genera preocupaciones, aunque la valoración global de la escala resulta pobre dado su reducido tamaño muestral. Al respecto del cuestionario DME, la valoración global de la escala resulta moderada y la valoración de la calidad del estudio revela riesgo de sesgo respecto del dominio prueba índice. Con relación al DT, es la herramienta que cuenta con un mayor número de estudios de validación. Su valoración global es moderada, y en cuatro de los cinco estudios no se estima riesgo de sesgo.

Por otra parte, localizamos 4 medidas cortas: dos para población paliativa (ESAS-r y ENP-E) y dos para población oncológica general (BSI-18 y HADS). La utilización del ESAS-r no estaría respaldada ni por la valoración global de la escala (pobre) ni por la valoración de la calidad de la evidencia, pues su único estudio de validación suscita preocupaciones tanto a nivel de riesgo de sesgo como de aplicabilidad. Con respecto al ENP-E, su valoración global resulta aceptable, mientras que la valoración de la calidad de la evidencia sugiere riesgo de sesgo incierto en la selección de pacientes y la prueba de referencia. En lo que concierne al BSI-18, la calidad de los estudios no plantea apenas preocupaciones sobre riesgo de sesgo y aplicabilidad, pero su valoración global resulta pobre, ya que no estima su precisión diagnóstica frente a una prueba de referencia. En cuanto a la HADS, su valoración global se considera moderada, mientras que la calidad de la evidencia de los estudios presenta riesgos de sesgo en dos de los tres estudios.

El compendio de todos los índices analizados indica que el Termómetro de distrés (con un punto de corte  $\geq 5$ ) destaca como herramienta de detección precoz del malestar emocional, por su brevedad (1 ítem), por haber sido validado en cinco estudios (de los cuales cuatro cuentan con bajo riesgo de sesgo) cuyas muestras provienen de población oncológica en diferentes estadios de la enfermedad, así como por contar con una valoración global de la escala moderada.

Para refinar el proceso de detección precoz del malestar emocional en pacientes con cáncer y superar los inconvenientes derivados de emplear una medida de un solo ítem, el uso del Termómetro de distrés podría emplearse en un primer estadio del propio cribado, y combinarse posteriormente con una escala de mayor longitud, lo que contribuiría a que la especificidad del cribado en su conjunto mejorase. La valoración conjunta de todos los índices analizados hace destacar a la escala HADS para este fin.

Cabe resaltar la importancia de emplear estas herramientas de detección precoz del malestar emocional dentro de un programa estructurado que asegure la atención psicosocial posterior a las y los pacientes que la necesiten, puesto que la evidencia no permite asegurar el efecto beneficioso del cribado de manera aislada.

## **Conclusiones**

- Para la elección de una u otra herramienta de detección precoz del malestar emocional en personas adultas con cáncer, deben considerarse aspectos como el objetivo de la detección precoz, la longitud de la escala, la fase de la enfermedad en la que se encuentran las y los pacientes, la valoración de riesgo de sesgos de los estudios y la valoración global de la escala.
- Considerando todo ello en su conjunto, destaca el Termómetro de distrés (con un punto de corte  $\geq 5$ ) como herramienta de detección precoz del malestar emocional por su brevedad (1 ítem), por haber sido validado en estudios con bajo riesgo de sesgo y cuyas muestras provienen de población oncológica en diferentes estadios de la enfermedad, así como por contar con una valoración global de la escala moderada. Esta medida de un solo ítem cuenta con una alta sensibilidad y una baja especificidad, por lo que podría considerarse como un primer estadio de cribado útil para descartar distrés clínico, y combinarse posteriormente con una escala de mayor longitud, lo que contribuiría a que la especificidad del cribado en su conjunto mejorase. Para este fin, destacaría la escala HADS.
- Cabe subrayar la relevancia de emplear las herramientas de detección precoz del malestar emocional en el marco de un programa de cribado estructurado y coordinado por un equipo multidisciplinar que cuente con la figura de un o una profesional de la psicología, así como con un protocolo de derivación que asegure la atención psicosocial especializada a las y los pacientes que la necesiten.
- Sería recomendable poner en marcha estudios que evalúen de manera completa la calidad psicométrica y la precisión diagnóstica de las herramientas de detección

precoz del malestar emocional, ya que el número de estudios es reducido y, en algunos casos, resultan incompletos al no incluir prueba de referencia, siendo particularmente necesario en el caso del BSI-18. Asimismo, sería necesario que futuras investigaciones presten mayor atención a cuestiones relacionadas con la implementación y la aceptabilidad de las herramientas por parte de los y las pacientes, sobre todo teniendo en cuenta la diversidad social y cultural, y las peculiaridades de grupos desfavorecidos o de población con discapacidad.

# Summary

**Introduction:** The model of care for people with cancer is being reformulated towards a holistic approach to health. This implies taking into account that their needs go beyond the physical aspects related to the disease and its treatment, including a wide range of emotional, interpersonal and/or social implications.

In this framework, emotional distress is defined by the US National Comprehensive Cancer Network (NCCN®) as an adverse multifactorial experience of a psychological, social, spiritual and/or physical nature that may interfere with the ability to effectively cope with cancer, its physical symptoms and its treatment. This distress extends along a spectrum ranging from feelings of vulnerability, sadness and fear to potentially disabling problems such as depression, anxiety, panic, social isolation, and existential or spiritual crisis.

Recent studies estimate that the prevalence of distress in the adult cancer population ranges between 25% and 52%. Aware of this prevalent reality, the Cancer Strategy of the National Health System (2021), within the strategic line on the care model, sets as one of its objectives the use in cancer care of a standardised tool for the early detection of emotional distress validated in Spanish.

**Objectives:** To provide an evaluation of the characteristics and psychometric properties of tools for the early detection of emotional distress in adult cancer patients validated in the Spanish context that can be used in the context of the National Health System (SNS).

**Methods:** Following a preliminary literature search of clinical practice guidelines (CPGs) and health technology assessment reports (conducted in October 2021), two specific search strategies were designed in the main biomedical databases: a first one for the location of clinical practice guidelines that address emotional distress in cancer patients (November 2021), and a second one with the objective of locating those primary studies of elaboration and validation of tools in the Spanish context, as well as those of cross-cultural adaptation and validation of tools to the Spanish context (in December 2021). This process was completed by creating alerts in the main databases (Medline, Embase and WoS) in order to retrieve all relevant studies that could be published during the drafting of the report. The alerts remained active until the date of publication of the report. After reading the titles and abstracts of the articles resulting from the search, a selection of studies was made according to the previously established inclusion and exclusion criteria. The selection of studies, data extraction and evidence synthesis process was peer-reviewed. The information was synthesised in evidence tables using the FLC 3.0 tool and adapted to the specific objectives of this report. The quality of the evidence of the studies was assessed using the QUADAS-2 tool for diagnostic efficacy studies. An overall assessment of each of the scales was also carried out, taking into account reliability, the type of criterion measure used, the sample sizes and the validity of each scale. The assessment of biases and applicability of the studies, as well as the overall assessment of the scales, was peer reviewed. The method for resolving disagreements in all phases was consensus.

In order to analyse the aspects related to the patients' perspective, the information on this subject was analysed both in the CPGs considered and in the primary studies included. In addition, this report was reviewed by 3 patient representatives from the Spanish Association Against Cancer (AECC).

**Results:** As a result of the guideline search strategy, 6 CPGs were taken into account to address the domains related to the description of the health problem, description and technical characteristics of the technology, as well as to analyse implementation considerations (organisational, ethical and social aspects and related to the patients' view). As results of the search for studies, 15 validations of tools for early detection of emotional distress in adult cancer patients were included. A total of 7 scales were evaluated: the Visual Analogic Scale (VAS) for anxiety and depression, the Distress Thermometer (DT), the Detection of Emotional Distress (DED) scale, the Brief Symptom Inventory-18 (BSI-18), the Edmonton Symptom Assessment Questionnaire - Revised (ESAS-r), the Hospital Anxiety and Depression Scale (HADS) and the Psychosocial and Spiritual Needs Evaluation scale (ENP-E). According to the length of the scale, 3 of the tools evaluated (VAS of anxiety and depression, DT and DED scale) are ultra-short (specifically between 1 or 2 items), and 4 of them (BSI-18, ESAS-r, HADS and ENP-E) are short (specifically between 10 and 18 items). Regarding the validation sample, 3 scales (DT, BSI-18 and HADS) were validated by three or more studies, with heterogeneous samples of cancer patients at different stages of the disease; while 4 scales (VAS for anxiety and depression, DED Scale, ESAS-r and ENP-E) were validated through a single study for each scale with oncological-palliative patients.

According to the QUADAS-2 tool for assessing the quality of diagnostic accuracy studies, there are no concerns about the risk of bias or the applicability of their results in the validation study of the VAS for Anxiety and Depression, four of the five validation studies of the DT, two of the three validation studies of the BSI-18 and one of the three validation studies of the HADS. The remaining studies raise concerns about either the risk of bias or the applicability of their results.

With respect to the psychometric properties analysed, the validation studies of the DT do not calculate either its internal consistency or its factorial structure, given that it is made up of a single item. In turn, the convergent validity of the test is assessed in all validation studies, reporting moderate positive correlations with other measures such as the BSI-18, the emotional problems included in the Problem List, or the HADS. The validation study of the VAS of anxiety and depression reports neither their internal consistency nor their factorial structure, understanding them as complementary measures, but not belonging to the same scale. However, it does analyse their convergent validity through the calculation of positive correlations of Global Emotional Distress (DEG) or score obtained from the sum of the VAS scores with the total HADS score, as well as between the anxiety VAS and the HADS anxiety subscale, and between the depression VAS and the HADS depression subscale. As regards the DED scale, its internal consistency ( $\alpha = 0.685$ ) is low, although this value is close to the minimum acceptable ( $\alpha = 0.7$ ), especially considering that the instrument consists of only two items. The validation study of this tool reports its convergent validity through the calculation of intraclass correlation coefficients (with the

DT, the HADS and its subscales, as well as with a clinical psychological assessment), which are positive. Validation studies of the BSI-18 report acceptable internal consistency ( $\alpha$  between 0.82 and 0.91), provide evidence of convergent validity and confirm a hierarchical structure with three first-order factors (anxiety, depression and somatization subscales) and a second-order factor as a measure of emotional distress (the Global Severity Index -GSI-). The ESAS-r also has an acceptable internal consistency ( $\alpha = 0.86$ ), but it does not analyse either the construct validity or the convergent validity of the scale. With reference to the HADS, the values provided by the validation studies on its internal consistency are acceptable ( $\alpha > 0.80$ ) and the underlying structure found in two of the studies is bifactorial, consisting of the anxiety and depression subscales. The other validation study finds a hierarchical solution with two first-order factors (anxiety and depression), and a second-order factor (total score) as a measure of emotional distress. In addition, one of the studies also provides evidence of convergent validity for each subscale. In the case of the ENP-E, the validation study reports acceptable internal consistency ( $\alpha = 0.75$ ) and an underlying four-factor structure, consisting of the subscales of emotional well-being, social support, spirituality and information. With regard to the convergent validity of the scale, the correlation coefficients calculated with the HADS and the DT were positive. The discriminative capacity of any of the early detection tools analysed proved to be good (with the area under the curve showing values above 0.80 in all cases). For the BSI-18 and ESAS-r, the discriminative capacity was not calculated due to lack of a reference test.

The overall assessment of each scale was based on a combined evaluation of its reliability, type of criterion measure used, sample sizes and validity of each scale (mean value of its sensitivity and specificity). Following this multi-criteria assessment, three scales were rated as poor (VAS of anxiety and depression, BSI-18 and ESAS-r), one scale was rated as acceptable (ENP-E) and three scales obtained a moderate rating (DT, DED scale and HADS).

The organisational impact of introducing tools for the early detection of emotional distress could be substantial, especially if it is introduced as part of a screening or systematic evaluation programme. Different CPGs recommend the creation of working teams to coordinate and direct its implementation, with a psychologist among their members. From an ethical point of view, the use of a tool for the early detection of emotional distress meets the requirement that the benefit outweighs the possible side effects. Other ethical considerations such as respect for autonomy and the right to privacy and confidentiality and equity of access must also be taken into account.

With regard to patients' views, hardly any information on acceptability, satisfaction or patients' perspectives has been identified in the CPGs retrieved or in the studies analysed.

**Discussion:** In general terms, the discriminative capacity of each of the tools analysed is good, although it should be noted that the results coincide with previous studies, which indicate that ultra-short measurements tend to have a lower diagnostic accuracy, as although they have adequate levels of sensitivity, the specificity values tend to be low to moderate. However, these measures represent an advantage in the healthcare setting as they are easier to implement in routine care.

In general, the use of one or the other tool may differ according to the context and type of patients with whom they are to be used. The quality of the evidence from the validation studies of each of the instruments, the number of validation studies supporting each tool, as well as the overall assessment of each scale, should also be taken into account when making such a choice.

Considering the results as a whole, we found three ultra-short measures, two (VAS for Anxiety and Depression, and the DED scale) for the palliative population, and one (Distress Thermometer) for the general oncology population. In the case of the VAS, the assessment of the quality of the only validation study does not raise concerns, although the overall assessment of the scale is poor given its small sample size. With regard to the DED scale, the overall assessment of the scale is moderate and the assessment of the quality of the study reveals a risk of bias with regard to the index test domain. The DT is the tool with the largest number of validation studies. Its overall assessment is moderate, and in four of the five studies no risk of bias is estimated.

We also found four short measures, two for the palliative population (ESAS-r and ENP-E) and two (BSI-18 and HADS) for the general oncology population. The use of the ESAS-r would not be supported either by the overall assessment of the scale (poor) or by the assessment of the quality of the evidence, as its single validation raises concerns both in terms of risk of bias and applicability. With respect to ENP-E, its overall assessment is acceptable, while the assessment of the quality of evidence suggests uncertain risk of bias in the selection of patients and reference test. Regarding the BSI-18, the quality of the studies raises few concerns about risk of bias and applicability, but its overall assessment is poor as it does not estimate its diagnostic accuracy against a reference test. In the case of the HADS, its overall assessment is considered moderate, while the quality of the evidence of the studies presents risks of bias in two of the studies.

The summary of all the indices analysed indicates that the distress thermometer (with a cut-off point  $\geq 5$ ) stands out as a tool for the early detection of emotional distress, due to its brevity (1 item), for having been validated in five studies (four with a low risk of bias) and whose samples come from the oncology population at different stages of the disease, as well as for having a moderate overall assessment of the scale.

To refine the process of early detection of emotional distress in cancer patients and overcome the disadvantages of using a single-item measure, the distress thermometer could be used at an early stage of the screening itself and then combined with a longer scale, which would help to improve the specificity of the screening as a whole. The combined assessment of all the indices analysed makes the HADS scale stand out for this purpose.

It is important to highlight the importance of using these tools for the early detection of emotional distress within a structured programme that ensures subsequent psychosocial care for those patients who need it, as the evidence does not allow us to ensure the beneficial effect of screening alone.

## Conclusions

- When choosing one or another tool for the early detection of emotional distress in adults with cancer, aspects such as the objective of early detection, the length of the scale, the stage of the disease in which the patients are, the assessment of the risk of bias of the studies and the overall assessment of the scale should be considered.
- All things considered, the distress thermometer (with a cut-off point  $\geq 5$ ) stands out as a tool for the early detection of emotional distress because of its brevity (1 item), because it has been validated in studies with a low risk of bias and whose samples come from cancer populations at different stages of the disease, and because it has a moderate overall assessment of the scale. This single-item measure has a high sensitivity and low specificity, so it could be considered as a useful first stage screening measure to rule out clinical distress, combined later with a longer scale, which would help to improve the specificity of the screening as a whole. For this purpose, the HADS scale stands out.
- It is important to stress the importance of using the tools for the early detection of emotional distress within the framework of a structured screening programme coordinated by a multidisciplinary team that includes a psychologist, as well as a referral protocol that ensures specialised psychosocial care for those patients who need it.
- It would be advisable to implement studies that comprehensively assess the psychometric quality and diagnostic accuracy of tools for the early detection of emotional distress, as the number of studies is small and, in some cases, incomplete because they do not include a reference test, which is particularly necessary in the case of the BSI-18. Future research would also need to give more attention to issues related to the implementation and acceptability of the tools by patients, especially taking into account social and cultural diversity, and the specificities of disadvantaged groups or people with disabilities.

# Justificación

Las necesidades de las y los pacientes con cáncer van más allá de los aspectos físicos relacionados con la enfermedad y su tratamiento, e incluyen un amplio abanico de implicaciones emocionales, interpersonales o sociales (1). Esto implica reformular el modelo de atención a la persona con cáncer hacia un enfoque integral de la salud (2). En este contexto, uno de los desarrollos recientes más importantes en psicooncología ha sido la elaboración y la validación de herramientas de detección precoz del malestar emocional (3).

La Red Nacional Integral sobre el Cáncer de EE.UU. (NCCN®) define el malestar o distrés emocional como una experiencia multifactorial desagradable de naturaleza psicológica, social, espiritual o física que puede interferir en la capacidad de afrontar de forma efectiva el cáncer, sus síntomas físicos y su tratamiento. Este malestar se extiende a lo largo de un continuo que comprende desde sentimientos de vulnerabilidad, tristeza y temor hasta problemas que pueden llegar a ser incapacitantes como la depresión, la ansiedad, el pánico, el aislamiento social y las crisis existenciales o espirituales (4).

Estudios recientes estiman que la prevalencia de distrés en población oncológica adulta oscila entre un 25 % y un 52 % (5-10). La gran variabilidad entre estudios con respecto a su metodología, los criterios establecidos para la definición de malestar psicológico, las herramientas de evaluación y el momento de llevarla a cabo pueden explicar dichas variaciones en la prevalencia (8).

A nivel nacional, la Estrategia en Cáncer del SNS (2021) (11), alineada con los objetivos del Plan Europeo de Lucha contra el Cáncer (12), se plantea como objetivo 28 «utilizar en la atención oncológica una herramienta estandarizada para la detección precoz del malestar emocional (*emotional distress*) validada al español que permita identificar, desde el momento del diagnóstico inicial, a aquellas personas con cáncer, sus familiares y las personas cuidadoras que pudieran necesitar atención psicológica, estableciendo un protocolo de derivación temprana para intervención psicológica en estos casos».

El propósito del presente informe es contribuir a la primera acción que propone este objetivo: la selección de las herramientas estandarizadas más adecuadas para la detección precoz del malestar emocional, a petición de la Dirección General de Salud Pública del Ministerio de Sanidad. Este informe se ha enmarcado en el Plan Anual de Trabajo de la Red Española de Agencias de Evaluación de Tecnologías Sanitarias y Prestaciones del SNS (RedETS).

# 1. Introducción

## 1.1. Descripción del problema de salud

### 1.1.1. Problema de salud diana

#### ¿Cómo se define el malestar emocional en personas adultas con cáncer?

El malestar emocional o distrés emocional ha sido definido por la NCCN® como una experiencia multifactorial desagradable de naturaleza psicológica, social, espiritual o física que puede interferir en la capacidad de afrontar de forma efectiva el cáncer, sus síntomas físicos y su tratamiento (13). Esta definición ha sido adoptada de manera general tanto en las guías de práctica clínica (GPC) como en la literatura científica sobre psicooncología.

El malestar emocional se extiende a lo largo de un continuo que abarca desde sentimientos de vulnerabilidad, tristeza y temor hasta problemas que pueden llegar a ser incapacitantes como la depresión, la ansiedad, el pánico, el aislamiento social y las crisis existenciales o espirituales (4). La siguiente tabla muestra los diferentes niveles de malestar en función de su presentación clínica (tabla 1).

**Tabla 1. Continuo del malestar emocional**

Bajo malestar	Malestar moderado o grave
– Se siente conectado/a con otras personas	– Se siente marginado/a y solo/a
– Cree que las cosas mejorarán	– Tiene un sentimiento de permanencia
– Puede disfrutar de recuerdos felices	– Siente culpa y arrepentimiento
– Tiene sentido de autoestima	– Siente autodesprecio
– El malestar tiene rachas	– El malestar es constante y continuo
– Está deseoso/a de cosas	– Está desesperanzado/a
– Todavía puede disfrutar de cosas	– No tiene interés
– Tiene ganas de vivir	– Presenta ideación suicida
– Tiene preocupaciones específicas	– Sufre ansiedad no focalizada
– Puede ver lo negativo y lo positivo	– Tiene una visión catastrófica
– Es capaz de tomar decisiones	– Es incapaz de comprometerse con el tratamiento de cáncer

**Fuente:** Tomada de *Pan-Canadian Practice Guideline: Screening, Assessment and Management of Psychosocial Distress, Depression and Anxiety in Adults with Cancer* (14).

## ¿Cuáles son los factores de riesgo relacionados con el malestar emocional en personas adultas con cáncer?

Las GPC consultadas ponen de manifiesto que las y los pacientes con mayor riesgo de sufrir malestar emocional moderado o grave son quienes presentan antecedentes de trastorno mental o trastorno por consumo de sustancias, así como deterioro cognitivo, enfermedades graves comórbidas, síntomas no controlados, barreras comunicativas y ciertos aspectos sociales como vivir solo o sola, tener hijos o hijas de corta edad, y tener antecedentes de trauma o abusos (físicos, sexuales, emocionales o verbales) (4). La evidencia científica señala también que el hecho de experimentar dolor continuo a lo largo de todo el curso de la enfermedad incrementa el riesgo de malestar emocional (15). En la siguiente tabla, se detallan todos los factores de riesgo identificados por las *NCCN® Clinical Practice Guidelines in Oncology* (4) y la *Pan-Canadian Practice Guideline* (14) (tabla 2).

**Tabla 2. Factores de riesgo de malestar emocional**

Historial médico y psicológico	Ataques de pánico, trastorno de ansiedad generalizada, historial de depresión o intento de suicidio, historial de trastornos del estado de ánimo, antecedentes de trastorno mental o trastorno por consumo de sustancias, antecedentes de trauma o abuso (físico, sexual, emocional o verbal).
Condiciones médicas	Enfermedades graves comórbidas, fase de tratamiento prolongada, deterioro cognitivo, intervenciones quirúrgicas, efectos secundarios del tratamiento, medicación actual asociada con la ansiedad o la depresión.
Puntos vulnerables en el curso de la enfermedad	Diagnóstico, control inadecuado del dolor o de otros síntomas, recidiva de la enfermedad, enfermedad avanzada o progresiva (metástasis), traslado a cuidados paliativos.
Tipo de cáncer	Cáncer de páncreas, de pulmón, de cabeza y de cuello.
Factores individuales	Edad más temprana, ser mujer, estilo de afrontamiento catastrofista o ansioso, cúmulo de acontecimientos vitales estresantes, cambios en las funciones y los roles asumidos.
Estado civil	Soltero/a, separado/a, divorciado/a o viudo/a.
Cuestiones sociales	Carencia de red de apoyo social, problemas económicos, acceso limitado a la atención médica, situación de inmigración, discriminación (racial, de género), pérdida de la vivienda, el refugio o el entorno estable, otras situaciones estresantes.
Condiciones familiares	Vivir solo/a, ser dependiente, tener hijos/as pequeños/as o hijos/as dependientes a cargo, conflictos familiares, conflictos con la persona cuidadora, mal funcionamiento marital o familiar.
Otros factores	Preocupaciones espirituales o religiosas, barreras comunicativas, mala comunicación con el equipo asistencial.

**Fuente:** Elaboración propia a partir de la compilación de los factores señalados por la NCCN® (4) y por la Asociación Canadiense de Oncología Psicosocial y la Asociación Canadiense contra el Cáncer (14).

## ¿Cuál es el curso del malestar emocional en personas adultas con cáncer?

Todas las personas adultas diagnosticadas de cáncer experimentan un cierto nivel de malestar asociado con este diagnóstico, los efectos de la enfermedad o su tratamiento, independientemente del estadio de la enfermedad en el que se encuentren (4). Para hacer frente a este malestar leve esperado, las personas utilizan una gran variedad de recursos, incluyendo sus propios recursos internos y el apoyo emocional de la familia y las amistades (14).

No obstante, algunos pacientes llegan a experimentar niveles de malestar moderado o grave, que pueden considerarse clínicamente significativos. Por este motivo, la identificación

y el tratamiento del malestar resulta de suma importancia a lo largo de todo el curso de la enfermedad y el tratamiento (4).

El malestar puede surgir de la anticipación del diagnóstico, como reacción al mismo, durante las fases tempranas del cáncer y el tratamiento (14) o de las sucesivas transiciones a través del curso de la enfermedad. De hecho, incluso cuando las respuestas psicológicas durante el tratamiento activo han sido satisfactorias, puede haber subgrupos de pacientes que todavía sean vulnerables a sufrir malestar *a posteriori*. (16).

En este sentido, existen periodos particularmente vulnerables al malestar emocional, entre los que se incluyen el momento del diagnóstico, el comienzo del tratamiento activo, las recurrencias, la transición a la supervivencia postratamiento —sobre todo cuando existe riesgo de recurrencia— y durante los cuidados paliativos y en la fase del final de la vida (14).

Si no se abordan las necesidades psicosociales de las y los pacientes con cáncer, independientemente de cuándo surjan, se convertirán en predictoras del aumento de efectos adversos y síntomas físicos (17-20), estrés y ansiedad posterior (21, 22), síntomas depresivos (23), deseo de morir pronto, tasas de mortalidad más elevadas y suicidio (14).

Por el contrario, el pronóstico de malestar puede mejorar si se interviene a tiempo. Los estudios muestran que las intervenciones psicosociales tempranas pueden ayudar a los y las pacientes a desarrollar estrategias de afrontamiento con las que responder al malestar emocional de manera precoz, lo que puede reducir el sufrimiento y el malestar a lo largo del proceso de enfermedad (14) y, en consecuencia, repercutir de forma positiva en su calidad de vida.

### **¿Con qué variables psicosociales está relacionado el malestar emocional en pacientes en edad adulta con cáncer?**

Los y las pacientes con malestar emocional pueden expresar miedo, preocupación, incertidumbre sobre la enfermedad y su futuro, tristeza, enfado, falta de sueño, falta de apetito, baja concentración, obsesión con pensamientos de enfermedad y muerte, y preocupación por sus roles y relaciones (14).

Por otra parte, se ha demostrado que sufrir malestar emocional es un factor de riesgo para una peor adherencia al tratamiento del cáncer (7, 24, 25). Asimismo, la presencia de malestar emocional en pacientes con cáncer puede llevar a una intensificación del dolor (15), una gravedad mayor de los síntomas físicos (14), un peor funcionamiento (14), una disminución en la calidad de vida del paciente (26), e incluso puede afectar negativamente a la supervivencia (27-30).

Cabe mencionar que, con independencia del momento y las circunstancias en las que la comorbilidad psiquiátrica surja vinculada al diagnóstico de cáncer, puede haber costes emocionales, interpersonales o económicos importantes para los y las pacientes (31).

### **¿Cuál es la prevalencia del malestar emocional en pacientes en edad adulta con cáncer y cuál es el impacto social del mismo?**

Los datos de prevalencia estimada de distrés emocional en población oncológica adulta a nivel internacional oscilan entre un 25 % y un 52 % (5, 6, 8-10, 32), datos similares a los hallados en los estudios realizados en población oncológica adulta en el territorio español, que oscilan entre el 26 % y el 37% (7, 34).

La gran variabilidad entre estudios con respecto a su metodología, los instrumentos de medida empleados, el punto en el curso de la enfermedad en el que se encuentren las y los pacientes, así como el entorno clínico donde se lleve a cabo la evaluación pueden explicar las variaciones halladas en los datos de prevalencia (8, 14), lo que dificulta inferir los datos de la prevalencia real (34).

La presencia de malestar emocional en pacientes con cáncer puede llevar a un aumento en los gastos de asistencia sanitaria (14, 35). Algunos estudios han demostrado que, sin importar las circunstancias y el momento en los que surge la comorbilidad psiquiátrica vinculada al diagnóstico de cáncer, puede haber consecuencias económicas para el sistema sanitario (31).

#### 1.1.2. Manejo actual del problema de salud

### **En la actualidad, ¿cuál es el abordaje terapéutico del malestar emocional en pacientes en edad adulta con cáncer?**

El manejo del malestar emocional abarca desde la prevención y la detección precoz hasta la derivación a servicios especializados en la atención psicológica necesaria para su tratamiento, así como la reevaluación en intervalos periódicos y el seguimiento.

La mayoría de las y los pacientes con cáncer son más capaces de adaptarse a llevar un diagnóstico de cáncer y los efectos de su tratamiento cuando tienen acceso a atención psicosocial efectiva que les ayuda a afrontar y resolver los problemas relacionados con la enfermedad (14). En este sentido, cada vez existe más evidencia de que la intervención precoz o psicosocial preventiva, como la información preliminar, el material de información o las intervenciones de autoayuda en línea, podrían ayudar a prevenir o reducir el malestar y a apoyar a los y las pacientes en su desarrollo de respuestas adaptativas tempranas en el curso del cáncer (14).

A nivel internacional, varias sociedades científicas, como la International Psycho-Oncology Society, y agrupaciones de profesionales expertos, como la European Partnership Action Against Cancer, coinciden en que se debería medir el distrés en la atención oncológica rutinaria como el sexto signo vital después de la temperatura, la presión arterial, el pulso, la frecuencia respiratoria y el dolor (36, 37). En la misma línea, la *European Guide on Quality Improvement in Comprehensive Cancer Control*, resultado del trabajo de la iniciativa Acción Conjunta Europea para el Control del Cáncer (CanCon), compuesta por especialistas en cáncer de 25 países y 126 organizaciones junto con la

Unión Europea, establece que los y las profesionales de la salud deberían llevar a cabo el *screening* o cribado periódico de malestar psicológico a lo largo de todo el proceso de la enfermedad, y a este debería seguirle una adecuada provisión de atención psicosocial (38). El malestar psicológico de los y las pacientes debería evaluarse de manera rutinaria en los momentos clave del curso de la enfermedad del paciente (14), pues se ha demostrado que la reevaluación sistemática en intervalos periódicos está asociada con una reducción de los síntomas de malestar a través del tiempo (14).

Diversas guías coinciden en que se debería cribar el malestar de todas y todos los pacientes en la primera consulta, los intervalos oportunos y cuando esté clínicamente indicado, sobre todo teniendo en cuenta los cambios en el curso de la enfermedad o el tratamiento —por ejemplo, el postratamiento, la remisión, la recurrencia, la progresión, las complicaciones con el tratamiento, la transición a cuidados paliativos y al final de la vida— (4, 14, 31). No obstante, las GPC ponen de manifiesto que a menudo las y los profesionales de la salud no reconocen los síntomas psicológicos y la atención social, y pueden infraestimar los beneficios del apoyo psicológico, de manera que no se les ofrece acceso a los servicios psicológicos (14).

El panel de la NCCN® recomienda que todas y todos los pacientes deberían cribarse mediante el Termómetro de malestar y la Lista de Problemas (4). En el caso de Canadá, las GPC de la Asociación Canadiense de Oncología Psicosocial y la Asociación Canadiense contra el Cáncer señalan que todas y todos los pacientes deberían pasar rutinariamente por un cribado de malestar y de los problemas que contribuyen a dicho malestar, empleando el cuestionario Edmonton Symptom Assessment System Scale Revised (ESAS-r), el Termómetro de distrés o malestar o el cuestionario Patient Health Questionnaire-2 (PHQ-2), junto con la lista Canadian Problem Checklist (14).

Según las recomendaciones de la NCCN®, en caso de que el malestar sea leve —lo que se denominaría «malestar esperado»—, lo abordaría el equipo de oncología (4). Por su parte, aquellos pacientes con un malestar moderado o grave requerirán de una derivación inmediata a un o una profesional de la salud mental, un asesor o una asesora de atención espiritual o a un trabajador o una trabajadora social, dependiendo del tipo de problemas identificados en la Lista de Problemas (4). Siguiendo las recomendaciones canadienses, aquellos/as pacientes con un malestar clínicamente significativo deberían pasar por una evaluación integral para detectar las fuentes, la naturaleza y el alcance del malestar, así como los factores de riesgo; y por una evaluación focalizada para identificar una posible sintomatología depresiva o ansiosa (14). La intervención psicológica adecuada dependerá de la naturaleza y la gravedad del problema psicológico de la persona, el diagnóstico, sus problemas psicológicos previos y la calidad del apoyo social disponible (14).

En el caso del contexto español, la última actualización de la Estrategia en Cáncer del SNS (11) reconoce que la calidad de vida, la atención psicológica y la evaluación de las necesidades de las y los pacientes que han sobrevivido al cáncer son retos no resueltos en nuestro SNS; de hecho, las asociaciones de pacientes y de voluntariado realizan tareas de sustitución que deberían ser parte del SNS. En ese mismo documento, se señala que este es el ámbito en el que nuestro sistema está más alejado de los países con mejores recursos

de apoyo —tanto psicológico como social— a las y los pacientes con cáncer (11). Por esta razón, algunas de las prioridades de actuación que indica la estrategia están en esta línea: (1) validar métodos breves y sencillos para detectar precozmente el malestar emocional (distrés), hacer una valoración del mismo y establecer el mejor abordaje para tratarlo; y (2) proporcionar atención psicológica y social a pacientes y familiares que la precisen en cualquier momento de la enfermedad, no limitándola al momento del alta y contando con los recursos necesarios, contribuyendo a disminuir el distrés psicológico y a mantener la calidad de vida de las y los pacientes y sus familias (11).

Tal y como evidencia el *Estudio sobre la atención psicológica en cáncer en el SNS en España* (2) realizado por la Asociación Española Contra el Cáncer (AECC), en la actualidad no se dispone de ningún tipo de registro oficial ni de información general sobre la forma en la que se proporcionan los servicios psicosociales, los recursos en el territorio y la organización de la atención a pacientes. Las autoras y los autores del estudio ponen de manifiesto que esta falta de datos dificulta entender realmente cómo se garantiza la atención psicológica especializada a pacientes con cáncer y familiares durante las diferentes etapas de la enfermedad en el territorio español (2). Complementariamente, ese informe analiza cómo las políticas autonómicas contemplan la provisión de atención psicológica a personas con cáncer y sus familiares por el SNS, e identifica tres escenarios principales: (1) atención psicológica especializada en cáncer provista por profesionales del SNS; (2) formalización de acuerdos de colaboración del sistema público de salud con entidades sin ánimo de lucro para la provisión de servicios de atención psicológica especializada en cáncer; o (3) inexistencia de acuerdos marco adecuados, lo que lleva a que las colaboraciones dependan de la proactividad de cada hospital en la gestión de convenios específicos.

### 1.1.3. Población diana

#### **¿Quiénes son los y las pacientes susceptibles de recibir el cribado de malestar emocional?**

A efectos de este informe de evaluación, las personas susceptibles de recibir el cribado o detección precoz del malestar emocional son aquellas y aquellos pacientes en edad adulta (mayores de 18 años) con diagnóstico de cualquier tipo de cáncer, sea cual sea el estadio o la fase de la enfermedad.

#### **¿Qué número potencial se estima de población oncológica adulta a nivel nacional?**

Para la estimación del tamaño de la población diana a nivel nacional, se han tenido en cuenta los datos aportados por la Estrategia en Cáncer del SNS (11). En dicho informe, se recoge la estimación de la prevalencia de cáncer en España a 31 de diciembre de 2020 realizada por la Red Española de Registros de Cáncer. El número estimado de casos prevalentes totales en España a 31 de diciembre de 2020 es de 2.265.152 (1.066.959 en hombres y 1.198.193 en mujeres) (11). Esto representa un total de 4611 casos por cada 100.000 hombres y de 4961 casos por cada 100.000 mujeres.

La población de España a 1 de enero de 2021 se cifra en 47.394.223 habitantes (INE, 2021), de forma que los casos prevalentes representan un 4,8 % de la población española. Teniendo en cuenta de manera conjunta estas cifras y los datos de prevalencia de malestar emocional moderado o grave en pacientes con cáncer reflejados por la literatura española (26-37 %) (7, 34), se podría hablar de entre 585.000 y 840.000 pacientes con malestar emocional moderado o grave en España.

## 1.2. Descripción y características técnicas de la tecnología a estudio

### 1.2.1. Características de la tecnología

#### **¿En qué consiste una herramienta para la detección precoz del malestar emocional en población oncológica adulta?**

La detección precoz consiste en la identificación de una enfermedad específica en una fase temprana del curso natural de la enfermedad (39). Teniendo en cuenta que, en ocasiones, no se identifica la enfermedad en sí misma sino un precursor, y que detección no es sinónimo de diagnóstico, para la confirmación de las sospechas diagnósticas se requerirá de test o procedimientos clínicos adicionales, propiamente diagnósticos (39). La detección precoz puede realizarse a través de la detección clínica temprana, el cribado oportunista o los programas de cribado poblacional.

En términos generales, las pruebas de cribado deben ser simples, rápidas, seguras, fáciles de realizar e interpretar, fiables, válidas, con alta sensibilidad —es decir, con una alta probabilidad de que una persona con malestar emocional clínicamente significativo sea identificada como tal por la prueba—, alto valor predictivo positivo —esto es, que exista una alta probabilidad de que aquellas personas identificadas con malestar emocional clínicamente significativo por la prueba, verdaderamente lo tengan—, bien aceptadas por profesionales y pacientes, y con una buena relación coste-efectividad (40).

Las herramientas de detección precoz en el caso de variables psicológicas consisten en test o escalas compuestas por varios ítems o preguntas que tienen por objeto medir algún aspecto psicológico (41). Asimismo, también pueden existir medidas de un solo ítem empleadas para la labor de la detección precoz.

Independientemente de su longitud, para poder emplear estas herramientas de detección precoz con garantías, es imprescindible que hayan sido validadas para el uso que se les pretende dar (40). En este tipo de pruebas resulta de especial relevancia la validez de constructo del instrumento. El término constructo se emplea en psicología para referirse a un concepto o una característica para cuya medición se diseña una prueba (42). A este respecto, algunas herramientas empleadas para la detección precoz del malestar emocional en población oncológica han optado por medir el malestar emocional como un constructo general, otras son herramientas multidimensionales que determinan los componentes que forman el malestar emocional, y otras son escalas unifactoriales relativas a uno de los dos

componentes que suelen formar el constructo de malestar emocional moderado o grave (ansiedad y depresión) (43).

La administración de estos instrumentos puede realizarse de manera heteroadministrada o autoadministrada.

Las herramientas de detección precoz han demostrado ser efectivas y fiables en la identificación del malestar y de las necesidades psicosociales de pacientes con cáncer (44-46). En ese sentido, las GPC hacen hincapié en que el *screening* de malestar emocional debe llevarse a cabo utilizando herramientas sencillas que estén validadas (47) y que cuenten con una sensibilidad y especificidad adecuadas, así como con puntos de corte establecidos para una rápida identificación del malestar moderado o grave (14, 31).

La revisión sistemática de Vodermaier *et al.* (48), publicada en 2009, sobre instrumentos en lengua inglesa de cribado de malestar emocional para pacientes oncológicos diferencia tres tipos de herramientas según su longitud, es decir, según el número de preguntas o ítems que las componen: ultracortas (entre 1 y 4 ítems), cortas (entre 5 y 20 ítems) y largas (entre 21 y 50 ítems), aunque explicita que el término herramienta de cribado suele emplearse para test particularmente breves. Las ventajas y desventajas de unas y otras herramientas en función de su longitud, según las personas autoras de la revisión (48), están recogidas en la siguiente tabla.

**Tabla 3. Ventajas y desventajas de las herramientas de cribado, según su longitud**

Medidas ultracortas (1-4 ítems)	Medidas cortas (5-20 ítems)	Medidas largas (21-50 ítems)
Excelente para usar en consultas concurrencias	Aceptable para usar en consultas concurrencias	Uso rutinario improbable a no ser que se automatice
La sensibilidad puede ser alta; la especificidad, de baja a moderada	Probablemente alta sensibilidad, especificidad de moderada a alta	La sensibilidad y la especificidad pueden ser altas
Puede evaluar un solo dominio	Puede evaluar múltiples dominios	Puede evaluar múltiples dominios
No adecuada para investigación	Puede ser adecuada para investigación, necesita ser probada	Excelente para investigación
Coste bajo	Cierto coste asociado a la puntuación (se puede minimizar a través de la automatización)	Puntuación potencialmente costosa (se puede minimizar a través de la automatización)

**Fuente:** Tomada de Vodermaier et al. (48).

### ¿Qué herramientas para la detección precoz del malestar emocional en **población oncológica adulta** recomiendan las guías de práctica clínica?

Con respecto a las GPC analizadas, aunque todas recomiendan llevar a cabo un cribado de malestar emocional en pacientes con cáncer, algunas no explicitan qué herramienta de cribado concreta debería utilizarse (49-51). A continuación, se incluye una tabla resumen de las GPC que se tuvieron en cuenta para elaborar este informe, y las herramientas recomendadas en cada caso.

**Tabla 4. Cuadro resumen de las GPC analizadas y las herramientas de detección precoz del malestar emocional en entorno clínico recomendadas por cada una de ellas**

Año	Organismo	Título de la GPC	Contexto	Herramientas específicas recomendadas
2022	Red Nacional Integral sobre el Cáncer de EE.UU. (NCCN®)	<i>NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines): Distress management. Version 1.2023</i>	EE.UU.	Distress Thermometer (DT) y Problem List
2021	Asociación Italiana de Oncología Médica	<i>Psychosocial care for adult cancer patients: Guidelines of the Italian Medical Oncology Association</i>	Italia	No se recomienda una herramienta específica
2018	Sociedad Europea de Oncología Médica (ESMO)	<i>Management of cancer pain in adult patients: ESMO Clinical Practice Guidelines</i>	Europa	No se recomienda una herramienta específica
2015	Asociación Canadiense contra el Cáncer y Asociación Canadiense de Oncología Psicosocial	<i>Pan-Canadian Practice Guideline: Screening, assessment and management of psychosocial distress, depression and anxiety in adults with cancer. Version 2.2015</i>	Canadá	Edmonton Symptom Assessment System Scale Revised (ESASr), DT o PHQ-2, junto con la Canadian Problem Checklist
2014	Sociedad Estadounidense de Oncología Clínica (ASCO)	<i>Screening, assessment, and care of anxiety and depressive symptoms in adults with cancer: An American Society of Clinical Oncology guideline adaptation</i>	EE.UU.	PHQ-2; Patient Health Questionnaire-9 (PHQ-9) y Generalized Anxiety Disorder7 (GAD7)
2004 (revisada en 2019)	Instituto Nacional para la Salud y la Excelencia Clínica del Reino Unido (NICE)	<i>Improving supportive and palliative care for adults with cancer</i>	Reino Unido	No se recomienda una herramienta específica

**Fuente:** Elaboración propia.

En el caso de la GPC de la NCCN®, se recomienda el uso del Distress Thermometer (DT) o Termómetro de malestar (52), una escala de valoración que va de 0 (No malestar) a 10 (Malestar extremo) (4). Esta herramienta buscó desestigmatizar el reporte del malestar presentándola en forma de termómetro (52). Con un solo ítem (Rodee el número —del 0 al 10— que describa mejor cuánto malestar ha experimentado durante la semana pasada, incluido el día de hoy), puede identificar el malestar experimentado por el o la paciente, sea cual sea su fuente, incluso aunque no esté directamente relacionado con el cáncer (4). La aplicación puede realizarse tanto de manera autoadministrada como heteroadministrada. El punto de corte propuesto por la NCCN® es el 4. Si la puntuación del paciente es menor a 4, su malestar sería considerado leve. Si la puntuación es mayor o igual a 4, el malestar sería considerado moderado o grave. En este último caso, sería necesario identificar el tipo de problemas a los que se enfrenta la o el paciente con la Lista de Problemas, que forma parte de la propia herramienta de *screening* propuesta por la guía (figura 1). Esta Lista de Problemas cubre 5 categorías distintas: problemas prácticos, familiares, emocionales, espirituales o religiosos y físicos, a través de 39 ítems. La herramienta en su conjunto ha sido validada en estudios con pacientes con diferentes tipos de cáncer, en diferentes entornos clínicos, culturas, países e idiomas (4). A continuación, se presenta la versión española de la herramienta propuesta por el NCCN Distress Management Panel.

**Figura 1. Versión española del Distress Thermometer**

**NCCN Guidelines, versión 2.2022  
Gestión del malestar**

**TERMÓMETRO DE MALESTAR NCCN**  
El malestar es una experiencia desagradable de naturaleza mental, física, social o espiritual. Puede afectar la forma de pensar, sentir o actuar. El malestar puede hacer que sea más difícil lidiar con el cáncer, sus síntomas o su tratamiento.

Instrucciones: Rodee el número (del 0 al 10) que describa mejor cuánto malestar ha experimentado durante la semana pasada, incluido el día de hoy.

**LISTA DE PROBLEMAS**  
¿Ha tenido alguna inquietud respecto a alguno de los siguientes puntos durante la semana pasada, incluido el día de hoy? (Marque todos los que procedan)

**Malestar extremo**

10  
9  
8  
7  
6  
5  
4  
3  
2  
1  
0

**Ningún malestar**

**Inquietudes físicas**

- Dolor
- Sueño
- Fatiga
- Consumo de tabaco
- Abuso de sustancias
- Memoria o concentración
- Salud sexual
- Cambios en la dieta
- Pérdida o alteración de capacidades físicas

**Inquietudes emocionales**

- Preocupación o ansiedad
- Tristeza o depresión
- Pérdida de interés o disfrute
- Duelo o pérdida
- Miedo
- Soledad
- Ira
- Cambios de aspecto
- Sensación de no valer nada o ser una carga

**Inquietudes sociales**

- Relación con el cónyuge o la pareja
- Relación con los hijos
- Relación con familiares
- Relación con amigos o compañeros de trabajo
- Comunicación con el equipo médico
- Capacidad de tener hijos

**Inquietudes de índole práctica**

- Cuidar de sí mismo
- Cuidar de los demás
- Trabajo
- Centro de estudios
- Cuidado del hogar
- Finanzas
- Seguros
- Transporte
- Cuidado de los niños
- Tener suficiente para comer
- Acceso a los medicamentos
- Decisiones acerca de tratamientos médicos

**Inquietudes de tipo espiritual o religiosa**

- El sentido de la vida
- Cambios de fe o creencias
- La muerte, morir o el más allá
- Conflicto entre las creencias y los tratamientos del cáncer
- Relación con lo sagrado
- Necesidades rituales o dietéticas

**Otras inquietudes:**

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Nota: Todas las recomendaciones son de categoría 2A, salvo que se indique lo contrario.  
Ensayos clínicos: NCCN considera que la mejor atención que puede recibir cualquier paciente con cáncer tiene lugar en el marco de un ensayo clínico. Se recomienda especialmente la participación en ensayos clínicos.

Versión 2.2022, 27/01/2022 © 2022 National Comprehensive Cancer Network® (NCCN®). Todos los derechos reservados. Queda prohibida cualquier reproducción de NCCN Guidelines® y las Instrucciones incluidas, en cualquier forma.

DIS-A

Fuente: Tomada de [https://www.nccn.org/professionals/physician\\_gls/pdf/distress-tool-spanish-eu.pdf](https://www.nccn.org/professionals/physician_gls/pdf/distress-tool-spanish-eu.pdf).

Con respecto a las propiedades psicométricas del DT, un metanálisis realizado en base a 42 estudios de distintos países: Australia (8), Países Bajos (5), Reino Unido (5), EE.UU. (4), Italia (3), China (3), Japón (2), uno multicéntrico en Italia, Portugal, Suiza y España (1), Nueva Zelanda (1), Dinamarca (1), Francia (1), Alemania (1), Islandia (1), Turquía (1), Irlanda (1), Corea (1), Suecia (1), Suiza (1) y España (1), con una muestra conjunta de más de 14 000 pacientes con cáncer, ha mostrado sus adecuados índices de sensibilidad (0,81; IC 95 % = 0,79–0,82) y especificidad (0,72; IC 95 % = 0,71–0,72) con el punto de corte 4 (53).

En el caso de la GPC canadiense (14), recomienda el uso de herramientas breves, con índices de sensibilidad y especificidad adecuados, y puntos de corte establecidos para el cribado de malestar emocional en personas adultas con cáncer, como el Edmonton Symptom Assessment Scale Revised (ESAS-r), el DT o el Patient Health Questionnaire-2 (PHQ-2).

El cuestionario Edmonton Symptom Assessment Scale fue desarrollado por Bruera *et al.* en 1991 (54) como una herramienta clínica para detectar la severidad sintomática de pacientes con cáncer avanzado en unidades de cuidados paliativos. A medida que su implementación se ha ido extendiendo, y con la revisión de la herramienta dando lugar al ESAS-r, su función ha cambiado y se ha convertido en una herramienta de cribado de malestar en pacientes con cáncer en entorno ambulatorio (55). Esta versión revisada

ha demostrado ser significativamente más sencilla de entender para los y las pacientes (55). El ESAS-r pregunta cómo se siente en ese momento el o la paciente con respecto a nueve síntomas básicos, así como un síntoma opcional que el paciente puede añadir. La valoración de cada ítem se hace con una escala tipo Likert de 11 puntos (de 0 a 10, donde 0 representa la ausencia del síntoma y 10 representa la mayor gravedad posible). La aplicación puede realizarse tanto de manera autoadministrada como heteroadministrada, dependiendo de la capacidad del paciente para poder informar de manera independiente (56). Las puntuaciones de cada ítem suelen interpretarse como ausencia del síntoma (0), leve (1-3), moderado (4-6) y grave (7-10) (57, 58).

**Figura 2. Cuestionario Edmonton Symptom Assessment System-Revised (ESAS-r)**



**APPENDIX A: Edmonton Symptom Assessment System Revised**

The ESAS-r and accompanying materials are available at: <http://palliative.org/tools.html>



**Edmonton Symptom Assessment System:  
(revised version) (ESAS-R)**

**Please circle the number that best describes how you feel NOW:**

No Pain	0	1	2	3	4	5	6	7	8	9	10	Worst Possible Pain
No Tiredness <i>(Tiredness = lack of energy)</i>	0	1	2	3	4	5	6	7	8	9	10	Worst Possible Tiredness
No Drowsiness <i>(Drowsiness = feeling sleepy)</i>	0	1	2	3	4	5	6	7	8	9	10	Worst Possible Drowsiness
No Nausea	0	1	2	3	4	5	6	7	8	9	10	Worst Possible Nausea
No Lack of Appetite	0	1	2	3	4	5	6	7	8	9	10	Worst Possible Lack of Appetite
No Shortness of Breath	0	1	2	3	4	5	6	7	8	9	10	Worst Possible Shortness of Breath
No Depression <i>(Depression = feeling sad)</i>	0	1	2	3	4	5	6	7	8	9	10	Worst Possible Depression
No Anxiety <i>(Anxiety = feeling nervous)</i>	0	1	2	3	4	5	6	7	8	9	10	Worst Possible Anxiety
Best Wellbeing <i>(Wellbeing = how you feel overall)</i>	0	1	2	3	4	5	6	7	8	9	10	Worst Possible Wellbeing
No _____ Other Problem <i>(for example constipation)</i>	0	1	2	3	4	5	6	7	8	9	10	Worst Possible _____

Patient's Name \_\_\_\_\_

Date \_\_\_\_\_ Time \_\_\_\_\_

Completed by (check one):

- Patient
- Family caregiver
- Health care professional caregiver
- Caregiver-assisted

**BODY DIAGRAM ON REVERSE SIDE**

Fuente: Tomada del manual de administración del ESAS-r (56).

Otra herramienta de *screening* de malestar emocional que propone la GPC canadiense (14) es el Patient Health Questionnaire-2 (PHQ-2) (59), formado por dos ítems que han sido extraídos de una herramienta más larga, el Patient Health Questionnaire-9 (PHQ-9) (60). El PHQ-2 pregunta por la frecuencia del estado de ánimo depresivo y la anhedonia durante las dos últimas semanas, puntuando cada ítem entre 0 (Nunca) y 3 (Casi cada día). La aplicación puede realizarse tanto de manera autoadministrada como heteroadministrada. En la siguiente figura, se puede consultar la versión española de dicha herramienta.

**Figura 3. Versión española del Patient Health Questionnaire-2 (PHQ-2)**

Durante las <u>últimas 2 semanas</u> ¿con qué frecuencia le ha molestado alguno de los siguientes problemas?	Nunca	Varios días	Más de la mitad de los días	
			Casi cada día	Cada día
a. Poco interés o alegría por hacer cosas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b. Sensación de estar decaído/a, deprimido/a o desesperanzado/a	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Fuente: Tomada de <https://www.phqscreener.com/select-screener>.

Con respecto a sus propiedades psicométricas, el estudio de validación original recomienda un punto de corte  $\geq 3$  para unos valores de sensibilidad de 0,83 y una especificidad de 0,90 (59). No obstante, un metanálisis posterior de estudios que evalúan las propiedades psicométricas del PHQ-2 sugiere que se podría usar un punto de corte  $\geq 2$  en situaciones de *screening* en las que se sabe que el colectivo al que se le aplica tiene un alto riesgo de depresión (61).

Según la herramienta de *screening* empleada, la GPC canadiense establece que, para aquellos y aquellas pacientes con puntuaciones iguales o superiores a 4 en el ESAS-r o en el DT, o iguales o superiores a 3 en el caso de emplear el PHQ-2, ha de completarse la evaluación con el uso de una herramienta específica que posibilite la identificación de las fuentes del malestar, su naturaleza y alcance (14), como, por ejemplo, la Canadian Problem Checklist. Esta herramienta fue desarrollada por el Grupo de Trabajo del Cancer Journey Action Group encargado de establecer el conjunto de datos mínimo necesario para el cribado de malestar, así como de desarrollar el kit de herramientas mínimas para la detección del malestar emocional (62, 63). La Canadian Problem Checklist está compuesta por 21 ítems que preguntan por problemas prácticos, emocionales, espirituales, sociales o familiares, físicos y de información. En la siguiente figura, puede consultarse esta herramienta.

**Figura 4. Canadian Problem Checklist**

Canadian Problem Checklist	
Please check all of the following items that have been a concern or problem for you in the past week including today:	
<b>Practical:</b> <input type="checkbox"/> Work/School <input type="checkbox"/> Finances <input type="checkbox"/> Getting to and from appointments <input type="checkbox"/> Accommodation	<b>Social/Family:</b> <input type="checkbox"/> Feeling a burden to others <input type="checkbox"/> Worry about family/friends <input type="checkbox"/> Feeling alone
<b>Emotional:</b> <input type="checkbox"/> Fears/Worries <input type="checkbox"/> Sadness <input type="checkbox"/> Frustration/Anger <input type="checkbox"/> Changes in appearance <input type="checkbox"/> Intimacy/Sexuality	<b>Informational:</b> <input type="checkbox"/> Understanding my illness and/or treatment <input type="checkbox"/> Talking with the health care team <input type="checkbox"/> Making treatment decisions <input type="checkbox"/> Knowing about available resources
<b>Spiritual:</b> <input type="checkbox"/> Meaning/Purpose of Life <input type="checkbox"/> Faith	<b>Physical:</b> <input type="checkbox"/> Concentration/Memory <input type="checkbox"/> Sleep <input type="checkbox"/> Weight

Fuente: Tomada de la *Guide to Implementing Screening for Distress, the 6th Vital Sign. Moving Towards Person-Centered Care* (62).

Por otra parte, la GPC de la Sociedad Estadounidense de Oncología Clínica (ASCO) recomienda emplear el PHQ-2 como herramienta de cribado para síntomas depresivos, con un punto de corte recomendado  $\geq 2$  (31). Para las y los pacientes con una puntuación igual o superior a 2, recomienda contestar al resto de ítems y completar el PHQ-9 (60). Además, recomienda emplear la escala Generalized Anxiety Disorder-7 (GAD-7) como herramienta de *screening* de síntomas ansiosos (64).

El PHQ-9, validado originalmente por Kroenke, Spitzer y Williams (60), consta de 9 ítems que preguntan sobre los 9 criterios en los que se basa el diagnóstico de los trastornos depresivos según el *Manual diagnóstico y estadístico de los trastornos mentales* (4ª ed.) (DSM-IV) (65). Se trata de un instrumento que puede aplicarse tanto de manera autoadministrada como heteroadministrada. La escala de respuesta de cada uno de los ítems va de 0 (Nunca) a 3 (Casi cada día), de manera que la puntuación total puede situarse entre 0 y 27 puntos. El PHQ-9 también cuenta con una pregunta final en la que se pide a los y las pacientes que informen de hasta qué punto estos problemas les han generado dificultades para realizar su trabajo, ocuparse de las tareas domésticas o relacionarse con las demás personas. Pese a no utilizarse en el cálculo de la puntuación total, este ítem puede ser útil en decisiones relativas al tratamiento, ya que está asociado a la gravedad de los síntomas psiquiátricos, así como a múltiples medidas de deterioro y calidad de vida relacionadas con la salud (66).

Con respecto a sus propiedades psicométricas, el estudio original de validación informa de un área bajo la curva de 0,95. Considerándolo una medida continua, las puntuaciones de 5, 10, 15 y 20 representan el umbral que marca los límites de un nivel leve, moderado, moderadamente grave y grave de síntomas depresivos (60). Asimismo,

una revisión sistemática sobre las propiedades psicométricas del PHQ-9 publicada en 2010 (67) puso de manifiesto sus adecuados valores de sensibilidad (entre 0,77 y 0,88) y especificidad (entre 0,88 y 0,94) para un punto de corte  $\geq 10$  en la detección de síntomas depresivos clínicamente significativos. En la siguiente figura, se puede consultar la versión española de dicha herramienta.

**Figura 5. Versión española del Patient Health Questionnaire-9 (PHQ-9)**

<b>CUESTIONARIO DE SALUD DEL PACIENTE-9 (PHQ-9)</b>				
<b>Durante las últimas 2 semanas ¿con qué frecuencia le ha molestado alguno de los siguientes problemas? (Para indicar su respuesta rodee el número con un círculo)</b>	<b>Nunca</b>	<b>Varios días</b>	<b>Más de la mitad de los días</b>	<b>Casi cada día</b>
1. Poco interés o alegría por hacer cosas	0	1	2	3
2. Sensación de estar decaído/a, deprimido/a o desesperanzado/a	0	1	2	3
3. Problemas para quedarse dormido/a, para seguir durmiendo o dormir demasiado	0	1	2	3
4. Sensación de cansancio o de tener poca energía	0	1	2	3
5. Poco apetito o comer demasiado	0	1	2	3
6. Sentirse mal consigo mismo/a; sentir que es un/a fracasado/a o que ha decepcionado a su familia o a sí mismo/a	0	1	2	3
7. Problemas para concentrarse en algo, como leer el periódico o ver la televisión	0	1	2	3
8. Moverse o hablar tan despacio que los demás pueden haberlo notado. O lo contrario: estar tan inquieto/a o agitado/a que se ha estado moviendo de un lado a otro más de lo habitual	0	1	2	3
9. Pensamientos de que estaría mejor muerto/a o de querer hacerse daño de algún modo	0	1	2	3

FOR OFFICE CODING   0   +        +        +         
=Total Score:       

---

**Si ha marcado alguno de los problemas de este cuestionario, ¿hasta qué punto estos problemas le han creado dificultades para hacer su trabajo, ocuparse de la casa o relacionarse con los demás?**

<b>Ninguna dificultad</b>	<b>Algunas dificultades</b>	<b>Muchas dificultades</b>	<b>Muchísimas dificultades</b>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Fuente: Tomada de <https://www.phqscreeners.com/select-screener>.

Por su parte, la escala Generalized Anxiety Disorder-7 (GAD-7) es una herramienta de *screening* de síntomas ansiosos desarrollada por Spitzer, Kroenke, Williams y Löwe (64). Se trata de un instrumento autoadministrado formado por siete ítems con una escala de respuesta tipo Likert que va de 0 (Nunca) a 3 (Casi todos los días). El punto de corte establecido es  $\geq 10$  para la detección de sintomatología moderada o severa (68). En la siguiente figura, se puede consultar la versión española de la herramienta.

**Figura 6. Versión española de la Generalized Anxiety Disorder-7 (GAD-7)**

<b>GAD-7</b>				
En los últimos 14 días, ¿con qué frecuencia le han supuesto una molestia los siguientes problemas?	Nunca	Varios días	Más de la mitad de los días	Casi cada día
<i>(Para indicar su respuesta rodee el número con un círculo)</i>				
1. Sentirse nervioso/a, angustiado/a o muy tenso/a	0	1	2	3
2. Ser incapaz de dejar de preocuparse o de controlar la preocupación	0	1	2	3
3. Preocuparse demasiado por diferentes cuestiones	0	1	2	3
4. Tener problemas para relajarse	0	1	2	3
5. Estar tan inquieto/a que le resulta difícil permanecer sentado/a	0	1	2	3
6. Enfadarse o irritarse con facilidad	0	1	2	3
7. Sentir miedo de que algo terrible pueda ocurrir	0	1	2	3
<b>(For office coding: Total Score T ___ = ___ + ___ + ___)</b>				

Fuente: Tomada de <https://www.phqscreeners.com/select-screener>.

Con relación a las propiedades psicométricas de la GAD-7, la validación original informa de una sensibilidad de 0,89 y una especificidad de 0,82, con el punto de corte 10 (64). En un metanálisis realizado en base a 9 estudios con una muestra conjunta de más de 4 600 participantes, se puso de relieve el punto de corte 8 como óptimo, dados sus adecuados índices de sensibilidad (0,83; IC 95 % = 0,71–0,91) y especificidad (0,84; IC 95 % = 0,70–0,92), aunque los puntos de corte del 7 al 10 contaban con índices de sensibilidad/especificidad similares (69).

## ¿Cuáles son los beneficios y los riesgos de la detección precoz del malestar emocional en pacientes en edad adulta con cáncer?

Desde un punto de vista general, tal y como señala Chamberlain (70), los beneficios de todo cribado consisten en una mejora del pronóstico de los casos detectados, un tratamiento menos radical que cura los casos detectados precozmente, un ahorro de recursos y una mayor tranquilidad en casos con resultado negativo. A su vez, la Organización Mundial de la Salud (71) apunta la necesidad de balancear los beneficios y los daños en cualquier programa de cribado. Entre los beneficios, señala el aumento de oportunidades de identificar condiciones de riesgo en un estadio precoz, reducir la severidad de la condición identificada, reducir la incidencia de la condición identificada y las muertes. Con respecto a los posibles daños, destaca el sobrediagnóstico, los falsos negativos, los falsos positivos y el desvío de recursos sanitarios para este fin (71).

De manera específica, el beneficio principal de la detección precoz del malestar emocional consiste en la posibilidad de abordaje y de intervención temprana de dicho malestar (4). Además, la detección precoz y el tratamiento del malestar emocional lleva a una mejor adherencia al tratamiento del cáncer, mejor comunicación profesional de la salud-paciente, menos llamadas y visitas a oncología, y evita el enfado de los y las pacientes, así como el desarrollo de trastornos mentales graves de ansiedad o depresión (4). No obstante, esto no se consigue únicamente con el uso de la herramienta de detección precoz, sino que a continuación son necesarias también las derivaciones y los tratamientos adecuados (4). El cribado y el tratamiento del malestar en cáncer beneficia a pacientes, familiares y personas cuidadoras, y al equipo médico, y ayuda a mejorar la eficiencia de las operaciones clínicas (4).

En la GPC italiana, se establece una recomendación fuerte de que el *screening* de malestar emocional debería ser considerado y diferentes programas de abordaje deberían ser implementados según el nivel de malestar emocional (49), lo que implica que considera que los beneficios de esta práctica son superiores a los riesgos. La GPC canadiense también establece recomendaciones fuertes con respecto a la realización del cribado y la intervención en el malestar emocional de los y las pacientes con cáncer (14), lo que, de nuevo, implica que considera que los beneficios superan claramente los daños y perjuicios.

Los riesgos de la detección precoz del malestar emocional no se han descrito en las GPC consultadas.

### 1.2.2. Regulación: licencias y autorizaciones

## ¿Existe algún requerimiento para la utilización de herramientas de detección precoz del malestar emocional en pacientes en edad adulta con cáncer?

Con respecto al uso de las herramientas, se debe contar con la autorización correspondiente en aquellos casos donde así lo hagan constar en su publicación. Por ejemplo, la NCCN® especifica que el uso del Termómetro de malestar necesita autorización expresa previa por escrito (4); mientras que en el caso del PHQ y la GAD-7, el *copyright* pertenece

a Pfizer Inc., que aclara en su web que no se precisa de permiso para la reproducción, la traducción, la presentación o la distribución de dichas herramientas (66, 68).

### 1.2.3. Utilización

#### **¿Cuál es nivel de utilización de herramientas de detección precoz del malestar emocional en pacientes en edad adulta con cáncer (en el SNS)?**

A nivel internacional, Canadá es el país donde la atención psicosocial forma parte del estándar de cuidado de los y las pacientes con cáncer, pues el malestar emocional está considerado la sexta constante vital y se revisa rutinariamente al igual que el resto de los signos vitales (72). El cribado de malestar emocional ha sido implementado en este país desde una perspectiva nacional y sus estrategias se han descrito en la literatura (73). También grupos de investigación de Italia, Francia, Países Bajos, EE.UU. o Japón han descrito los resultados de sus esfuerzos hacia la implementación del cribado de malestar emocional (74-78). Sin embargo, se desconoce con exactitud cuál es el nivel de implementación de las herramientas de detección precoz en los diferentes sistemas nacionales de salud.

En el caso de España, a pesar de las altas implicaciones de la morbilidad psicológica en la población oncológica, esta sigue siendo un área clínica frecuentemente subestimada (11). Con relación a ello, el *Estudio sobre la atención psicológica en cáncer en el SNS en España* recomienda la medición del estrés o malestar emocional como el sexto signo vital, incorporando el valor de las mediciones en la historia clínica (2). No obstante, como ya se mencionó, no se dispone de ningún tipo de registro oficial o información general sobre la forma en que se proporcionan los servicios psicosociales, los recursos en el territorio y la organización de la atención a pacientes, lo cual dificulta conocer cómo se garantiza la atención psicológica especializada a pacientes con cáncer durante las diferentes etapas de la enfermedad (2).

A este respecto, la Estrategia en Cáncer del SNS aprobada en 2021 subraya como uno de sus objetivos utilizar en la atención oncológica una herramienta estandarizada para la detección precoz del malestar emocional (11). Para alcanzar este objetivo, propone desarrollar las siguientes acciones: (1) seleccionar la/s herramienta/s estandarizada/s más adecuada/s para la detección precoz del malestar emocional; (2) introducir la/s herramienta/s en la historia clínica; (3) elaborar un protocolo de derivación a atención psicológica especializada; y (4) realizar un estudio del impacto psicológico del cáncer y sus tratamientos en el marco del SNS (11). El hecho de que parte de las acciones que se deben desarrollar según la Estrategia en Cáncer del SNS vigente sean la selección y la introducción de la/s herramienta/s más adecuada/s en la historia clínica hace pensar que el nivel de utilización actual de herramientas de detección precoz del malestar emocional en pacientes en edad adulta con cáncer es irregular e insuficiente en el marco de nuestro SNS.

#### 1.2.4. Requerimientos de la técnica

##### **¿Qué personal sanitario puede realizar la detección precoz del malestar emocional en pacientes en edad adulta con cáncer y en qué contexto de asistencia sanitaria?**

Aunque en todas las GPC consultadas se recomienda realizar un cribado de malestar emocional en pacientes con cáncer, en algunas de ellas no se especifica qué profesionales han de encargarse de llevarlo a cabo (49, 50).

Según el panel de la NCCN®, el Termómetro de malestar puede facilitarlo el personal de recepción o auxiliares de clínica (4). Siguiendo la GPC de la ASCO y la GPC canadiense, cualquier profesional de la salud puede encargarse de cribar la presencia de malestar emocional en pacientes con cáncer (14, 31). En el caso de la GPC del NICE, se establece que el cribado de malestar emocional debiera realizarlo el personal que haya sido designado y entrenado adecuadamente para tal función (personal de enfermería especialista, trabajadores o trabajadoras sociales y médicos o médicas de familia) en cada equipo de oncología y cuidados paliativos especializados (51).

Sean quienes fueren las y los profesionales que brinden las herramientas de detección precoz del malestar emocional a los y las pacientes, es necesario que el personal sanitario encargado de la interpretación de los resultados conozca el instrumento, su sistema de puntuación, el punto de corte establecido y el protocolo que debe seguir según la puntuación obtenida por el o la paciente.

En cuanto al contexto de asistencia sanitaria donde podría efectuarse el *screening*, algunas de las GPC consultadas no especifican esta información (49, 50). Por su parte, el panel de la NCCN® establece que el DT puede administrarse en entornos ambulatorios y hospitalarios (4). La GPC canadiense y la de la ASCO indican que las recomendaciones formuladas en dicha guía son aplicables a los y las profesionales sanitarias en los distintos entornos asistenciales (14, 31) y la GPC del NICE recomienda que el *screening* de malestar emocional en pacientes con cáncer se realice en los servicios de oncología y cuidados paliativos (51).

##### **¿Qué materiales son necesarios para emplear las herramientas de detección precoz del malestar emocional en población oncológica adulta y en qué tipo de instalaciones se puede efectuar?**

En general, a la hora de utilizar una herramienta de *screening* de malestar emocional, puede optarse o bien por lápiz y papel, para una administración tradicional, o bien por un dispositivo informático (ordenador, tableta o app), en caso de que se prefiera una administración digital.

Las ventajas y los inconvenientes de utilizar una u otra administración fueron analizados por el Screening for Distress Toolkit Working Group de Canadá (62) (ver tabla 5). Aunque el grupo no llegó a formular ninguna recomendación formal, sí apuntó que, cuando es posible el *screening* digital, ofrece más ventajas que el cribado con papel y lápiz.

**Tabla 5. Ventajas e inconvenientes del cribado con lápiz y papel versus cribado digital**

Cribado con lápiz y papel		Cribado digital	
Ventajas	Inconvenientes	Ventajas	Inconvenientes
Los y las pacientes están acostumbrados a completar formularios en papel	Hay que digitalizar los datos para la evaluación, el análisis de tendencias, etc.	La investigación indica que la mayoría de las y los pacientes considera que los cuestionarios digitales son fáciles de usar	Puede ser necesario ayudar a las y los pacientes que no se sientan cómodos con el uso de la tecnología, al menos la primera vez
No hay que invertir en software ni en hardware	Puede dar lugar a montones de papel	La información de tendencias puede imprimirse inmediatamente y los informes pueden trasladarse a la historia médica del paciente	Se requiere soporte informático; se puede necesitar un nuevo espacio si se usa un mostrador para ello
No se requieren nuevos espacios	Puede disminuir la integración de los datos	Resulta sencillo realizar una evaluación del programa de <i>screening</i>	Pueden necesitarse copias disponibles en lápiz y papel para las y los pacientes que estén demasiado enfermos para completarlo en un ordenador

**Fuente:** Tomada de la *Guide to Implementing Screening for Distress, the 6th Vital Sign. Moving Towards Person-Centered Care. Background, Recommendations, and Implementation* (62).

En lo que concierne a las instalaciones donde se puede llevar a cabo el *screening* de malestar emocional, algunas guías indican que las evaluaciones y las intervenciones psicológicas deberían efectuarse en instalaciones que sean tranquilas, cómodas y que confieran privacidad (51).

## 2. Alcance y objetivo

### 2.1. Objetivo

El objetivo de este informe es proporcionar una evaluación de las características y las propiedades psicométricas de las herramientas para la detección precoz del malestar emocional en población adulta oncológica validadas en el contexto español que puedan utilizarse en el contexto del SNS.

### 2.2. Alcance

Este informe tiene como población objetivo a pacientes en edad adulta con cáncer. Después de explorar cuáles son las herramientas recomendadas por las GPC para la detección precoz del malestar emocional en esta población, se analizarán las características y las propiedades psicométricas de aquellas validadas en el contexto español y que, por tanto, puedan utilizarse en el contexto del SNS. Se tendrán en cuenta las validaciones realizadas en cualquier contexto de asistencia sanitaria.

El presente informe está dirigido a la Dirección General de Salud Pública y al personal clínico y sanitario del SNS involucrado en la asistencia a la población oncológica.

A continuación, se describe la pregunta de investigación transformada en formato PIRDS (población objetivo, pruebas índice, pruebas de referencia, condición de interés y diseño de estudios) (tabla 6).

**Tabla 6. Pregunta PIRDS (población objetivo, prueba índice, prueba de referencia, condición de interés y diseño de estudios)**

Descripción	Alcance
Población objetivo	Pacientes en edad adulta con cáncer (Código CIE-11: O2).
Pruebas índice	Cualquier herramienta de detección precoz del malestar emocional validada en pacientes en edad adulta con cáncer en el contexto español.
Pruebas de referencia	Juicio clínico, otras herramientas diagnósticas o de cribado previamente validadas [por ejemplo, la Hospital Anxiety and Depression Scale (HADS) o el Brief Symptom Inventory-18 (BSI-18)].
Condición de interés	Malestar emocional.
Diseño de estudios	Estudios transversales o casos-control de elaboración y validación de herramientas en el contexto español. Estudios transversales o casos-control de adaptación transcultural y validación de herramientas al contexto español. Revisiones sistemáticas (RS), metanálisis.

## 3. Método

### 3.1. Metodología de elaboración del informe

#### 3.1.1. Estrategia de búsqueda bibliográfica

Tras una búsqueda preliminar de guías de práctica clínica e informes de evaluación de tecnologías sanitarias (ETS), se diseñaron dos estrategias de búsqueda específicas: una estrategia de búsqueda para la localización de guías de práctica clínica que aborden el malestar emocional en población oncológica, y otra estrategia para dar respuesta a la pregunta PIRDS desarrollada anteriormente, con el objetivo de localizar aquellos estudios de elaboración y validación de herramientas en el contexto español, así como aquellos de adaptación transcultural y validación de herramientas al contexto español.

#### Bases de datos consultadas

Las estrategias diseñadas se adaptaron a las bases de datos especificadas en la tabla 7.

Se emplearon bases de datos y motores de GPC, bases de datos generales y específicas. Además, se realizó una búsqueda en bases de datos de tesis doctorales e informes de ETS con el fin de localizar estudios de elaboración y validación de herramientas en el contexto español.

**Tabla 7. Bases de datos bibliográficas de literatura biomédica consultadas**

Tipos	Bases de datos
Bases de datos y motores de búsqueda de guías de práctica clínica	<ul style="list-style-type: none"><li>- Tripdatabase</li><li>- G-I-N</li><li>- Guíasalud</li><li>- NICE</li></ul>
Bases de datos generales	<ul style="list-style-type: none"><li>- Medline (Pubmed/Ovid)</li><li>- Embase (Ovid)</li><li>- Web of Science</li></ul>
Bases de datos específicas de psicología, enfermería y cuidados	<ul style="list-style-type: none"><li>- CINAHL (EBSCO)</li><li>- PsycINFO (Ovid)</li></ul>
Bases de datos de literatura gris e informes de Evaluación de Tecnologías Sanitarias (ETS)	<ul style="list-style-type: none"><li>- Teseo</li><li>- Dialnet</li><li>- International HTA Database</li></ul>

Fuente: Elaboración propia.

En la estrategia de búsqueda se emplearon diferentes descriptores y términos libres: *emotional stress, distress syndrome, psychological distress, psychometry, thermometer, rating scale, neoplasms, cancer*, entre otros. Los términos libres se buscaron en los campos de título, el resumen y las palabras clave. En el anexo A se pueden consultar las estrategias íntegras que se utilizaron en cada una de las bases de datos.

Todo este proceso se completó mediante la creación de alertas en las principales bases de datos (Medline, Embase y WoS) con el fin de recuperar todos los estudios relevantes que pudiesen ser publicados durante la redacción del informe. Las alertas se mantuvieron activas hasta la fecha de edición del informe (diciembre de 2022).

El resultado de la búsqueda se volcó en un gestor de referencias bibliográficas (EndNote 20.2) para eliminar los duplicados y facilitar la gestión documental.

Tras la lectura de los títulos y los resúmenes de los artículos derivados de la búsqueda, se procedió a seleccionar los estudios a texto completo, con revisión manual de la bibliografía referida en los mismos, para acceder a todos aquellos estudios que pudieran ser relevantes. Este procedimiento se llevó a cabo por pares.

### **Fecha de ejecución de la/s búsqueda/s. Actualizaciones de la búsqueda inicial**

La búsqueda de guías de práctica clínica se efectuó en noviembre de 2021, y la búsqueda de estudios de elaboración y validación de herramientas en el contexto español se realizó en diciembre de 2021. Se recibieron alertas mensuales hasta diciembre de 2022.

#### **3.1.2. Criterios de selección de estudios**

La selección de guías tuvo en cuenta que el documento se ajustase a la definición de GPC —«conjunto de recomendaciones basadas en una revisión sistemática de la evidencia y en la evaluación de los riesgos y los beneficios de las diferentes alternativas, con el objetivo de optimizar la atención sanitaria a los pacientes» (79)—, los criterios temporales —guías publicadas en los últimos 10 años— y de contenido —que abordasen la condición a estudio—.

La selección de los estudios primarios se realizó por pares en base a los criterios de inclusión/exclusión establecidos según la pregunta PIRDS. Se planteó que los desacuerdos fuesen resueltos por consenso. Los criterios de selección fueron:

- **Diseño del estudio:** revisiones sistemáticas, metanálisis, estudios transversales y casoscontrol de elaboración o validación de herramientas. Se excluyeron las revisiones narrativas, los estudios observacionales retrospectivos o no consecutivos y los casos clínicos.
- **Tipo de publicación:** artículos originales de investigación, tesis doctorales. Se excluyeron las cartas al director, los artículos de opinión, los editoriales y las comunicaciones a congresos.
- **Población a estudio:** pacientes en edad adulta con cáncer. Se excluyó la población infantil y adolescente.
- **Tamaño muestral:** sin limitación.
- **Pruebas índice:** herramientas de detección precoz del malestar emocional validadas en el contexto español. Se excluyeron las herramientas validadas en otros países.

- Pruebas de referencia: juicio clínico, otras herramientas diagnósticas o de cribado previamente validadas.
- Idioma de la publicación: español e inglés. Se excluyó cualquier otro idioma distinto a los citados.

### 3.1.3. Extracción de datos y síntesis de la evidencia

Tanto la extracción de datos como el proceso de síntesis de la evidencia se realizó por pares. Una autora extrajo de forma sistemática las principales características de cada uno de los estudios de validación seleccionados, pormenorizando los índices y los coeficientes calculados en tablas de evidencia. Estas tablas se elaboraron específicamente para este informe a partir de las fichas de lectura crítica cubiertas a través de la herramienta FLC 3.0 (80) sobre cada uno de los estudios analizados. La otra autora confirmó la pertinencia y la exhaustividad de estas. El método de resolución de desacuerdos fue el consenso. Las tablas han sido incluidas en el apartado de anexos del informe (anexo B).

En el apartado de resultados del informe, se compendió en una tabla resumen la información relativa a las características de las herramientas de cribado incluidas —el número de ítems, la escala de respuesta, el marco temporal de referencia y el rango de puntuación total—; y, en otra, se recapituló la información contextual sobre los estudios de validación española seleccionados, indicando el instrumento original y, si es el caso, la adaptación española utilizada, la muestra de validación empleada y el ámbito sanitario donde fue validada.

También se sintetizaron las propiedades psicométricas de cada una de las validaciones de las escalas, informando específicamente sobre su fiabilidad y validez. Las tablas resumen recogen información sobre los índices que han sido empleados en al menos la mitad de los estudios, de manera que facilitan la comparación de propiedades entre estudios y entre escalas.

La consistencia interna se refiere al grado en el que los ítems de una escala se correlacionan entre ellos (81), es decir, la magnitud en la que miden el mismo constructo (82). La consistencia interna mide, valga la redundancia, el grado de consistencia de las respuestas respecto al dominio psicológico medido, por lo que, si los ítems que componen una escala teóricamente miden el mismo constructo, la escala debe mostrar un alto grado de homogeneidad (81, 82). Este es el método más utilizado para medir la fiabilidad de los instrumentos, siendo el coeficiente de alfa de Cronbach ( $\alpha$ ) el índice que más se emplea para medirla en escalas politómicas.

Con respecto a la validez, en la actualidad la Asociación Estadounidense de Psicología (APA) la define como el grado en el que la evidencia y la teoría respaldan las interpretaciones de las puntuaciones de una prueba para los usos propuestos de dicha prueba (42). En este sentido, para que una prueba sea considerada útil para un propósito determinado, se necesitan evidencias que respalden que la prueba mide lo que pretende medir (83). Desde esta teoría contemporánea de validez, no se habla de tipos de validez,

sino de fuentes de evidencia de validez: evidencias de validez basada en el contenido de la prueba, los procesos de respuesta, la estructura interna, la relación con otras variables y las consecuencias de la aplicación de las pruebas (42).

No obstante, los estudios de validación analizados emplean el enfoque clásico, aportando principalmente información sobre validez convergente, que se refiere al grado de relación entre distintos procedimientos que miden el mismo constructo (84), y sobre validez de constructo, que se entiende como el grado en el que un instrumento mide la dimensión evaluativa para la que fue diseñada, es decir, determina la relación del instrumento con la teoría y la conceptualización teórica (85), determinando el número de dimensiones subyacentes que componen tal constructo y que están siendo medidas por la escala. El análisis factorial es, por excelencia, la técnica utilizada para este tipo de validación (86).

Asimismo, se reunió la información sobre la precisión diagnóstica de las herramientas de cribado en tablas resumen, que compendian los resultados para cada escala analizada. En concreto, se ha informado sobre el área bajo la curva, la sensibilidad, la especificidad, el valor predictivo positivo, el valor predictivo negativo, la utilidad clínica positiva, la utilidad clínica negativa y el valor global de cada prueba. Las tablas han sido incluidas en el apartado de resultados del informe.

Complementariamente, se realizó una valoración global de cada una de las escalas siguiendo el procedimiento llevado a cabo por Vodermaier *et al.* (48) en su revisión sistemática sobre instrumentos en lengua inglesa de cribado de malestar emocional para población oncológica, publicada en 2009. Los criterios de análisis fueron: la fiabilidad, el tipo de medida criterio empleada y la validez de cada escala —entendiendo por ello el valor medio de su sensibilidad y especificidad—. Estos criterios pueden recibir una valoración alta, moderada o baja. La combinación de estas tres calificaciones da como resultado una valoración global que puede ser excelente, buena, moderada, aceptable o pobre (48). La valoración global se consideraría pobre si alguno de los tres criterios se evalúa como bajo, si la fiabilidad no ha sido reportada (en escalas con 3 o más ítems), si no se efectúa el análisis de la curva ROC (Receiver Operating Characteristic), o si el número de participantes es menor de 100 cuando se emplea como medida criterio una escala autoadministrada, o menor de 50 cuando se emplean como medida criterio las entrevistas clínicas estructuradas. La valoración global sería excelente en el caso de que los tres criterios hayan recibido una valoración alta; sería buena si dos de los criterios reciben valoraciones altas y uno, moderada; y sería calificada como moderada en caso de que un criterio recibiese una valoración alta, y dos recibiesen una moderada. Si los tres criterios reciben una valoración moderada, la valoración global sería aceptable.

#### 3.1.4. Valoración de la calidad de la evidencia

Se planteó que la calidad de la evidencia se valoraría empleando las escalas específicas para cada tipo de estudio, en base a las recomendaciones de la *Guía para la elaboración y adaptación de informes rápidos de evaluación de tecnologías sanitarias* desarrollada dentro de la línea de trabajos metodológicos de la RedETS (87):

- Revisiones sistemáticas y metanálisis: herramienta AMSTAR-2.
- Estudios transversales y de casos-controles de elaboración o validación de herramientas de cribado de malestar emocional: herramienta QUADAS-2.
- Estudios transversales y de casos-controles de adaptación transcultural de herramientas: directrices para la traducción y adaptación de los test (2ª edición) de la International Commission Test.

Sin embargo, una vez seleccionados los estudios pertinentes, todos resultaron ser estudios transversales de validación, por lo que la herramienta empleada para valorar la evidencia en todos los casos fue la QUADAS-2 (88, 89), que puede consultarse en el anexo D.

El proceso de evaluación de la calidad de los estudios, la valoración sobre los sesgos y la aplicabilidad se realizó por pares. En concreto, una autora realizó la evaluación de cada uno de los estudios, mientras que la otra autora confirmó su pertinencia y adecuación. El método de resolución de desacuerdos establecido fue el consenso.

### 3.1.5. Metodología para la incorporación de la visión de los y las pacientes

Para examinar los aspectos relacionados con la visión de los y las pacientes se analizó la información al respecto tanto en las GPC consideradas como en los estudios primarios incluidos.

Además, este informe fue revisado por 3 representantes de pacientes de la Asociación Española Contra el Cáncer (AECC).

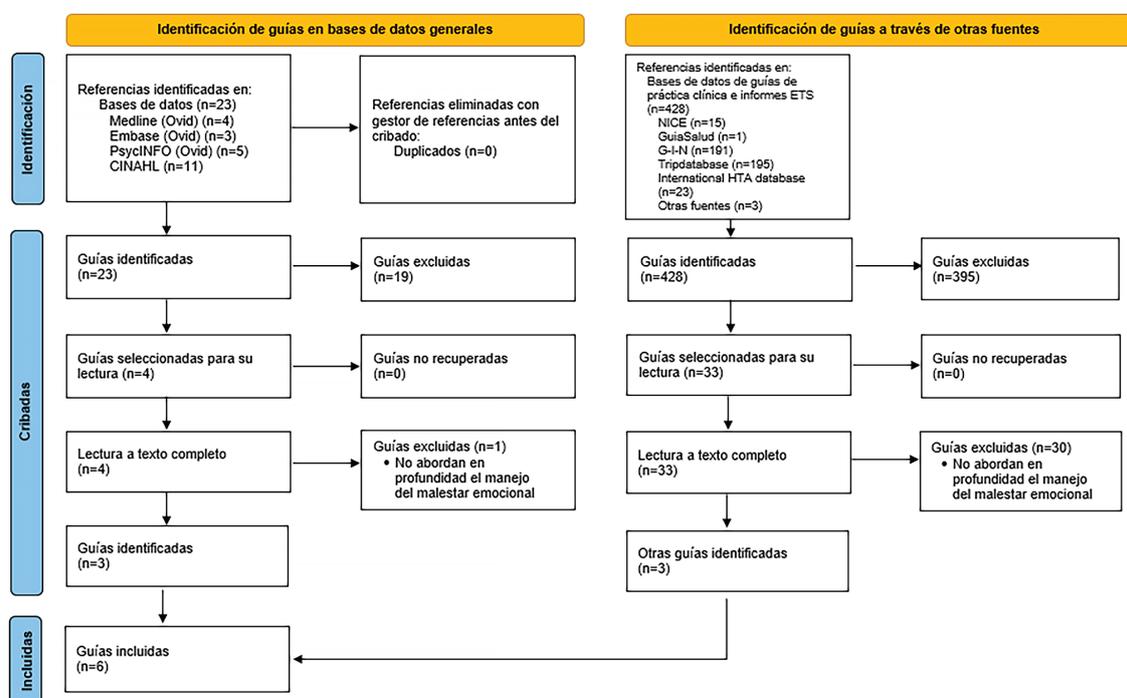
## 4. Resultados

### 4.1. Descripción de la evidencia disponible

#### 4.1.1. Resultados de la búsqueda

La búsqueda de guías de práctica clínica arrojó un total de 23 referencias en las bases de datos bibliográficas de literatura biomédica y 428 referencias en las bases de datos y los motores de búsqueda de guías de práctica clínica. La siguiente figura detalla el diagrama de flujo de resultados de dicha búsqueda (figura 7). El nivel de acuerdo entre pares fue del 100 %.

**Figura 7. Diagrama de flujo de los resultados de la búsqueda inicial de guías de práctica clínica**



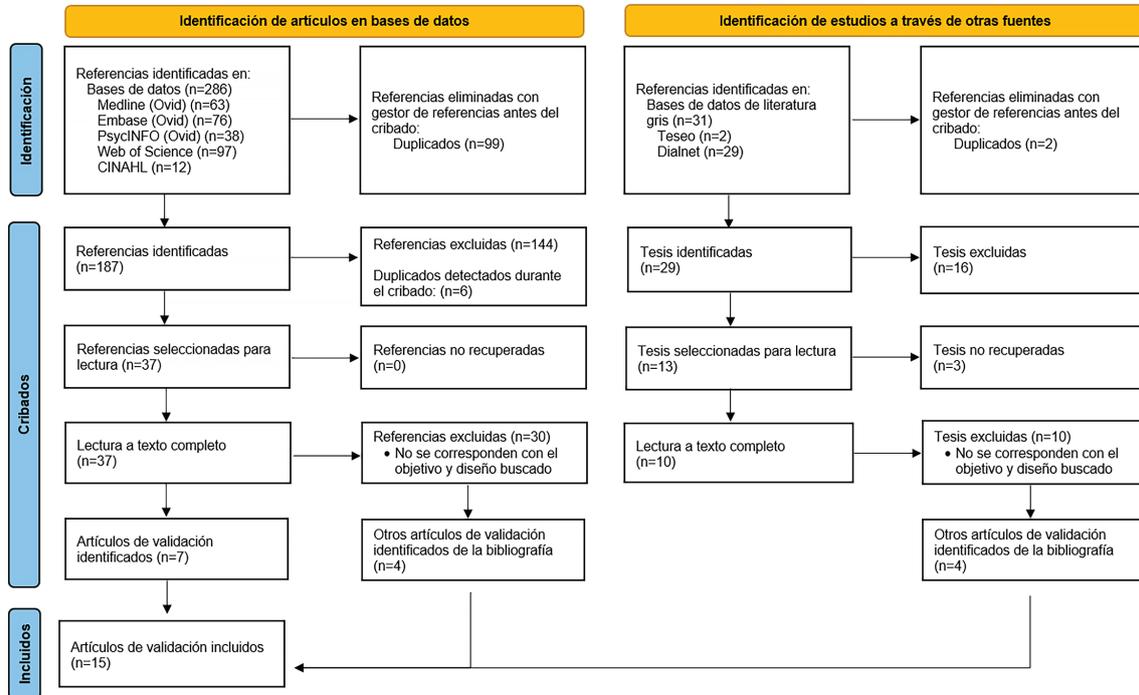
**Fuente:** Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. PubMed PMID: 33782057.

La búsqueda de estudios de elaboración y validación de herramientas en el contexto español produjo 286 referencias en las bases de datos bibliográficas habituales, y 31 en las bases de datos de tesis. En la siguiente figura, se muestra el diagrama de flujo del proceso de inclusión de estudios (figura 8).

Tras la lectura del título y el resumen de un total de 187 referencias, se seleccionaron 37 estudios para su lectura a texto completo, pues consideraban el malestar emocional como una de las variables de estudio. De ellos, finalmente se incluyeron 7 estudios de validación por cumplir los criterios especificados en el apartado anterior. Asimismo, se realizó una revisión manual de la bibliografía de los 30 estudios excluidos para asegurar la inclusión de

otros estudios de validación relevantes no recuperados en la búsqueda bibliográfica. Así, se recuperaron 4 estudios más de validación.

**Figura 8. Diagrama de flujo de los resultados de la búsqueda bibliográfica de estudios primarios**



**Fuente:** Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. PubMed PMID: 33782057.

Por otra parte, se llevó a cabo la lectura del título y el resumen de 29 tesis doctorales. De ellas, se seleccionaron 13 para su lectura a texto completo, pues contemplaban el malestar emocional como una de las variables a estudio. Pudieron recuperarse un total de 10. Dichas tesis fueron excluidas, pero, mediante la revisión manual de su bibliografía se accedió a 4 estudios adicionales de validación.

En resumen, los estudios primarios que finalmente se incluyeron suman un total de 15. El nivel de acuerdo inicial entre las autoras fue del 87 %, y los desacuerdos existentes se resolvieron por consenso. Las referencias y las tesis excluidas tras su lectura a texto completo, así como las causas de exclusión, se detallan en el anexo C.

#### 4.1.2. Descripción de las guías de práctica clínica seleccionadas

Se seleccionaron un total de 6 GPC, desarrolladas en distintos contextos. En concreto, una de ellas se desarrolló en Canadá, por parte de la Asociación Canadiense contra el Cáncer y la Asociación Canadiense de Oncología Psicosocial (14); otra se llevó a cabo en el marco del Instituto Nacional para la Salud y la Excelencia Clínica del Reino Unido (NICE) (51), y otra contó con el respaldo de la Asociación Italiana de Oncología Médica (49). Asimismo, dos fueron desarrolladas en Estados Unidos —una promovida por la Sociedad Estadounidense de Oncología Clínica (ASCO) (31) y la otra por la Red Nacional Integral

sobre el Cáncer de EE.UU. (4)—, y otra en el marco europeo, amparada por la Sociedad Europea de Oncología Médica (ESMO) (50).

Sus resultados se tuvieron en cuenta para abordar los dominios relativos a la descripción del problema de salud, la descripción y las características técnicas de la tecnología, así como para analizar las consideraciones de implementación —los aspectos organizativos, los éticos y sociales, y los relacionados con la visión de los y las pacientes—.

#### 4.1.3. Descripción y calidad de los estudios seleccionados

Todos los estudios seleccionados son estudios prospectivos transversales, realizados en España con población oncológica. Las muestras de los estudios están compuestas por pacientes en diferentes fases de la enfermedad —con diagnóstico reciente (90, 91), en fase de tratamiento (35, 92-98), en fase de seguimiento (94, 96-99), en fase de supervivencia (99) o en fase avanzada o de cuidados paliativos (100-103)— y con diagnósticos de diferentes tipos de cáncer primario —principalmente de mama (90-99, 101, 103), ginecológico (92, 96, 97, 99), genitourinario (92, 95-99), gastrointestinal (91, 93, 94, 99-101, 103), de pulmón (91, 94, 95, 99, 101, 103), melanoma (98, 99), de cabeza y cuello (101) o hematológico (35)—. Doce de los estudios seleccionados se han realizado en un sola comunidad autónoma —7 en la Comunidad Valenciana (35, 90, 92, 96-99), 3 en Cataluña (95, 101, 102), 1 en Cantabria (91) y 1 en Navarra (100)—, mientras que 2 fueron estudios multicéntricos llevados a cabo en diferentes comunidades autónomas (93, 94) y uno no aclara este punto (103).

Las herramientas evaluadas en estos estudios son: el Inventario Breve de Síntomas18 (BSI-18) (n=3 estudios); el Cuestionario de evaluación de síntomas de Edmonton, versión revisada (ESASr) (n=1 estudio); las Escalas Visuales Analógicas (EVA) de ansiedad y depresión (n=1 estudio); la Escala Hospitalaria de Ansiedad y Depresión (HADS) (n=3 estudios); el Termómetro de distrés (DT) (n=5 estudios); el Cuestionario de Detección de Malestar Emocional (DME) (n=1 estudio) y el Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida (ENP-E) (n=1 estudio).

Descripción de las escalas utilizadas en los estudios seleccionados

El Brief Symptom Inventory-18 (BSI-18) o Inventario Breve de Síntomas18 (104) es una herramienta compuesta por 18 ítems que preguntan sobre la frecuencia de diferentes síntomas con una escala tipo Likert de 0 a 4 puntos. Estos ítems se agrupan en tres subescalas (somatización, depresión y ansiedad), y en una puntuación total llamada Índice Global de Severidad (GSI). El marco temporal de referencia es la semana previa.

El Edmonton Symptom Assessment System-revised (ESAS-r) o Cuestionario de evaluación de síntomas de Edmonton, versión revisada (55) es un instrumento que evalúa la presencia de diez síntomas, mediante una escala numérica visual que va desde 0 a 10. A mayores, cuenta con un ítem abierto donde el o la paciente puede incluir algún otro problema y evaluarlo del mismo modo. El marco temporal de referencia es el momento actual (ahora).

Las Escalas Visuales Analógicas (EVA) de ansiedad y depresión evaluadas por Martínez *et al.* (101) se utilizan de manera conjunta como medida de malestar emocional, dando lugar a una puntuación total llamada Distrés Emocional Global.

La Hospital Anxiety and Depression Scale (HADS) o Escala Hospitalaria de Ansiedad y Depresión está compuesta por 14 ítems que preguntan sobre la frecuencia o la intensidad de distintos síntomas, medidos con una escala de respuesta tipo Likert de 0 a 3 puntos (105). Los ítems se agrupan en dos subescalas (ansiedad y depresión), y en una puntuación total como medida de malestar emocional. El marco temporal de referencia es la semana previa.

Por su parte, el Distress Thermometer (DT) o Termómetro de distrés (52) es una escala analógica visual con puntuaciones que van desde 0 (Ausencia de malestar) a 10 (Malestar extremo). El marco temporal de referencia es la semana previa. En algunos estudios se estudia su combinación con el Impact Thermometer (IT) o Termómetro de impacto, que emplea una escala visual de 11 puntos para determinar la interferencia en la vida cotidiana del malestar emocional, siendo 0 la ausencia de interferencia y 10 la interferencia extrema (106).

El Cuestionario de Detección de Malestar Emocional (DME) consta de dos partes (107). La primera parte está formada por dos preguntas en formato de escala visual numérica de 0 a 10, en las que se evalúa el estado de ánimo y la percepción de afrontamiento de la situación, junto con otra pregunta que registra la presencia o ausencia de preocupaciones; la segunda parte consiste en una observación, por parte de un o una profesional de la salud, de la presencia de signos externos de malestar emocional. El marco temporal de referencia es el momento actual.

Por último, el Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida (ENP-E) consta de tres secciones (103). La sección A está compuesta por 13 ítems que evalúan la intensidad con la que las y los pacientes han tenido ciertos sentimientos, pensamientos o conductas recientemente. La escala de respuesta es tipo Likert de 5 puntos. La sección B explora la intensidad de las preocupaciones de los y las pacientes sobre ciertos aspectos de su vida (financiero, familiar, emocional, espiritual, físico y otros), con una escala tipo Likert de 5 puntos. La sección C indaga sobre la presencia de seis signos externos de malestar emocional, valorados de 0 a 10 por parte del entrevistador o la entrevistadora.

En la tabla 8 se sintetizan las características de todas las escalas analizadas.

#### Descripción de los estudios

Se localizaron tres estudios que analizan las propiedades psicométricas del **Inventario Breve de Síntomas18 (BSI-18)** (90, 92, 93). Aunque en estos estudios no se usa una prueba de referencia para examinar la precisión diagnóstica del instrumento, se decidió incluirlos puesto que esta herramienta se emplea como prueba de referencia en otros estudios para la validación de escalas que también se analizan. En el estudio de Galdón *et al.* (90), publicado en 2008, la muestra estuvo formada por 175 pacientes mujeres que habían sido

**Tabla 8. Características de las herramientas de detección precoz analizadas**

Instrumento original	N.º de ítems / Escala de respuesta	Marco temporal / Rango de puntuación total
Brief Symptom Inventory-18 (BSI-18; Derogatis, 2000) Inventario Breve de Síntomas-18 (BSI-18)	Dieciocho ítems sobre la frecuencia del síntoma, medidos con una escala de respuesta tipo Likert de 0 a 4 puntos. Los ítems se agrupan en tres subescalas (somatización, depresión y ansiedad), y en una puntuación total llamada Índice Global de Severidad (GSI).	El marco de referencia temporal es la semana previa (es una medida de estado). El rango de puntuación total va de 0 a 24 puntos en cada subescala (somatización, depresión y ansiedad), y de 0 a 72 en la puntuación total (GSI).
Edmonton Symptom Assessment System-revised (ESAS-r; Watanabe, Nekolaichuk, Beaumont, Johnson, Myers y Strasser, 2011) Cuestionario de evaluación de síntomas de Edmonton, versión revisada	Evalúa la presencia de diez síntomas, mediante una escala numérica visual que va desde 0 a 10. Además, hay un ítem abierto donde el o la paciente puede incluir algún otro problema y evaluarlo del mismo modo.	El marco de referencia temporal es el momento actual (ahora). El rango de puntuación total va de 0 a 10 puntos para cada síntoma de manera independiente.
Visual Analogic Scales (VAS) for anxiety and depression Escalas Visuales Analógicas (EVA) de ansiedad y depresión	Dos preguntas en formato de escala analógica visual con puntuaciones desde 0 a 10. La suma de las puntuaciones de ambas escalas da como resultado la puntuación de Distrés Emocional Global (DEG)	No se explicita cuál es el marco de referencia temporal. El rango de puntuación de cada escala va de 0 a 10 puntos, y el rango de puntuación total va de 0 a 20.
Hospital Anxiety and Depression Scale (HADS; Zigmond y Snaith, 1983) Escala Hospitalaria de Ansiedad y Depresión	Catorce ítems sobre la frecuencia o la intensidad del síntoma, medidas con una escala de respuesta tipo Likert de 0 a 3 puntos. Los ítems se agrupan en dos subescalas: ansiedad (7 ítems) y depresión (7 ítems). También se puede calcular una puntuación total como medida de malestar emocional.	El marco de referencia temporal es la semana previa (es una medida de estado). El rango de puntuación total va de 0 a 21 puntos en cada subescala (42 en total).
Distress Thermometer (DT; Roth, Kornblith, Batel-Copel, Peabody, Scher y Holland, 1998) Termómetro de distrés	Escala analógica visual con puntuaciones desde 0 (Ausencia de malestar) a 10 (Malestar extremo).	El marco de referencia temporal es la semana previa (es una medida de estado). El rango de puntuación total va de 0 a 10 puntos.
Cuestionario de Detección de Malestar Emocional (DME; Maté, Mateo, Bayés, Bernaus, Casas, González-Barboteo, Limonero, López-Postigo, Sirgo y Viel, 2009)	El cuestionario consta de dos partes: a) Dos preguntas en formato de escala visual numérica de 0 a 10, en las que se evalúa el estado de ánimo y la percepción de afrontamiento de la situación, y otra pregunta que registra la presencia o la ausencia de preocupaciones. b) La segunda parte consiste en una observación, por parte de la o el profesional de salud, de la presencia de signos externos de malestar emocional.	El marco de referencia temporal es el momento actual. El DME otorga una puntuación total (de 0 a 20) formada por la suma de las puntuaciones de las preguntas en formato de escala visual numérica. Las preguntas sobre preocupaciones o signos externos de malestar emocional no se valoran cuantitativamente.
Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida (ENP-E; Mateo-Ortega, Limonero, Maté-Méndez, Beas, González-Barboteo, Barbero, Ela y Gómez-Batiste, 2019)	Sección A: 13 ítems (contesta el o la paciente). Escala de respuesta tipo Likert de 5 puntos. Sección B: 6 ítems (contesta el o la paciente). Escala de respuesta tipo Likert de 5 puntos. Sección C: 6 ítems (detectados por quien entrevista). Escala de respuesta de 0 a 10.	El marco temporal de referencia es recientemente. Se entiende que el rango de la puntuación total de la sección A será de 5 a 65, pero no se explicita. La sección B no se tiene en cuenta cuantitativamente. No se aclara cómo se calcula la puntuación total de la sección C.
<b>Fuente:</b> Elaboración propia.		

diagnosticadas de cáncer de mama recientemente y acudían a la Fundación Instituto Valenciano de Oncología (FIVO, Valencia). En el estudio de Martínez *et al.* (92), publicado en 2019, se incluyó una muestra de 1183 pacientes oncológicos en edad adulta que acuden a la FIVO (Valencia) para recibir asistencia sanitaria en cualquiera de sus departamentos médicos. Un año más tarde, Calderon *et al.* (93) publicaron un estudio realizado con 877 pacientes con cáncer no avanzado que acudían a alguno de los 15 departamentos de oncología médica participantes (provenientes de 10 comunidades autónomas).

También se localizó un estudio donde se analizan las propiedades psicométricas del **Cuestionario de evaluación de síntomas de Edmonton, versión revisada (ESAS-r)** (100). En él no se emplea una prueba de referencia para estudiar la precisión diagnóstica del instrumento, y la muestra estuvo formada por 66 pacientes con cáncer avanzado que acuden o bien al departamento de oncología del Hospital de la Universidad de Navarra (España) o al del Instituto Nacional del Cáncer (Guatemala). A pesar de que la muestra no es en su totalidad de España y de que no cuenta con una prueba de referencia, decidió incluirse el estudio en el cuerpo de evidencia porque esta herramienta es una de las recomendadas por las GPC analizadas previamente (14).

Por otra parte, se seleccionó un estudio que evalúa la capacidad de cribado de malestar emocional mediante las **Escalas Visuales Analógicas (EVA) de ansiedad y depresión** (101), dando lugar a una puntuación total llamada Distrés Emocional Global. La prueba de referencia utilizada fue la Escala Hospitalaria de Ansiedad y Depresión (HADS) y la muestra la conformaron 53 pacientes con diagnóstico de cáncer avanzado o terminal que se atendieron por primera vez en la consulta externa del Servicio de Cuidados Paliativos del Instituto Catalán de Oncología (ICO) (101).

Asimismo, se incluyeron 3 estudios de validación de la **Escala Hospitalaria de Ansiedad y Depresión (HADS)** (94-96). López-Roig *et al.* (94) validaron la escala HADS con una muestra de 484 pacientes oncológicos que acuden a consultas externas, hospital de día, planta hospitalaria o consulta de psicología de siete hospitales y centros de diferentes puntos de España (Comunidad Valenciana, Castilla y León, Comunidad de Madrid y Canarias). Las pruebas de referencia empleadas fueron el Inventario de Ansiedad Estado-Rasgo (subescala de ansiedad-estado), STAI-S, (108) para la subescala de ansiedad, y la Escala de depresión de Zung, SDS, (109) para la subescala de depresión. Por su parte, Costa *et al.* (95) realizaron su estudio con una muestra de 234 pacientes oncológicos ambulatorios que acudían al hospital de día o a las consultas en las unidades funcionales del Hospital Duran y Reynals (Barcelona), perteneciente al ICO. La prueba de referencia empleada fue la Entrevista clínica estructurada para los trastornos del eje I del DSMIV, SCID-I; (110). En el caso de Martínez *et al.* (96), la muestra de su estudio estuvo formada por 892 pacientes oncológicos que visitaban la FIVO (Valencia) por cualquier razón (diagnóstico, tratamiento, seguimiento), y la prueba de referencia utilizada fue la versión española del BSI-18 (104).

Con respecto al **Termómetro de distrés (DT)**, se localizaron un total de 5 estudios que llevaron a cabo su validación en población oncológica española (35, 91, 97-99) —analizándose complementariamente en dos de ellos la combinación del DT con el

Termómetro de impacto (IT) (35, 98)—. El primer estudio publicado de validación del DT en población española es el de Martínez *et al.* (97) en 2003, con una muestra de 962 pacientes con cáncer, que acuden a la FIVO (Valencia) para recibir tratamiento en cualquiera de sus departamentos médicos. La prueba de referencia empleada fue el BSI-18 (104). Un año más tarde, Priede *et al.* (91) publicaron su estudio de validación con una muestra de 135 pacientes con un primer diagnóstico reciente de cáncer que acuden al Hospital Universitario Marqués de Valdecilla (Santander). Las pruebas de referencia utilizadas fueron la versión española de la escala HADS (111) y una combinación de la clasificación realizada por la HADS junto con una entrevista diagnóstica semiestructurada basada en los criterios diagnósticos DSM-IV. En 2015, se publicaron dos artículos que estudian la validación del DT y su posible mejora al combinarlo con el IT (35, 98): el de Martínez *et al.* (98) con una muestra de 385 pacientes oncológicos en edad adulta que visitaban algún departamento de la FIVO (Valencia), y el de Andreu *et al.* (35) con una muestra de 81 pacientes oncológicos en edad adulta hospitalizados con neoplasias hematológicas en el Hospital Universitario la Fe de Valencia. En ambos estudios, la prueba de referencia utilizada fue el BSI-18. Posteriormente, el DT también fue validado por Andreu *et al.* (99) con una muestra de 236 pacientes en edad adulta supervivientes tras recibir tratamiento de cáncer de inicio en la edad adulta que acuden a la FIVO (Valencia) para citas médicas de seguimiento. La prueba de referencia empleada también fue el BSI-18.

Además, se recuperó un estudio de validación del **Cuestionario de Detección de Malestar Emocional (DME)** (102). La muestra de validación estuvo formada por 105 pacientes oncológico-paliativos hospitalizados en unidades de cuidados paliativos de cinco hospitales de Cataluña (102). Como pruebas de referencia emplearon: (1) una entrevista clínica semiestructurada realizada por un psicólogo o una psicóloga; (2) el Termómetro de distrés (DT) y (3) la versión española de la escala HADS validada por Tejero *et al.* (112).

Para terminar, se incluyó un estudio de validación del **Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida (ENP-E)** (103). La muestra de validación estuvo formada por 150 pacientes oncológico-paliativos hospitalizados en unidades de cuidados paliativos de 19 hospitales participantes en el estudio (103). De ese conjunto de 150 pacientes, 21 conformaron la muestra del estudio piloto, y 129, la muestra del estudio de validación. Como pruebas de referencia emplearon una combinación de cuatro medidas: (1) la versión española de la escala HADS validada por Tejero *et al.* (112); (2) el ítem sobre calidad de vida del Cuestionario de Calidad de Vida Core 15 Palliative de la Organización Europea para la Investigación y el Tratamiento del Cáncer [EORTC QLQ-C15-PAL (113)]; (3) el Termómetro de distrés (DT) y (4) la percepción subjetiva del paso del tiempo (114, 115).

En la tabla 9 se compendian las descripciones de todos los estudios analizados.

Las características pormenorizadas de cada uno de los estudios incluidos y su calidad se muestran en las tablas de evidencia (anexo B).

**Tabla 9. Descripción contextual de los estudios de validación seleccionados**

Instrumento original	Estudio de validación española en población oncológica	Versión española utilizada	Muestra validación española / Ámbito sanitario
Brief Symptom Inventory-18 (BSI-18; Derogatis, 2000).	Galdón, Durá, Andreu, Ferrando, Murgui, Pérez e Ibañez, 2008	Traducción española publicada por el autor original (Derogatis, 2000).	175 pacientes mujeres que habían sido recientemente diagnosticadas de cáncer de mama y acuden a la FIVO (Valencia, España). Edad media = 52,7 años (dt = 10,12). Rango de edad entre 27 y 70 años.
	Martínez, Conchado, Andreu y Galdón, 2019	Versión española publicada por el autor original (Derogatis, 2013).	1183 pacientes oncológicos en edad adulta que acuden a la FIVO (Valencia, España) para recibir asistencia sanitaria en cualquiera de sus departamentos médicos. 65 % mujeres. Edad media = 57,71 años (dt = 13,09). Rango de edad entre 19 y 94 años.
Inventario Breve de Síntomas-18 (BSI-18)	Calderon, Ferrando, Lorenzo-Seva, Hernández, Oporto-Alonso y Jiménez-Fonseca, 2020	No consta.	877 pacientes con cáncer confirmado histológicamente, no avanzado y tratado quirúrgicamente, que acuden a alguno de los 15 departamentos de oncología médica de 10 comunidades autónomas participantes en el estudio, y para quienes las directrices clínicas internacionales consideran que el tratamiento adyuvante es una opción. 59,7 % mujeres. Edad media = 36 años (dt = 13,4). Rango de edad entre 25 y 84 años.
Edmonton Symptom Assessment System-revised (ESAS-r; Watanabe, Nekolaichuk, Beaumont, Johnson, Myers y Strasser, 2011).  Cuestionario de evaluación de síntomas de Edmonton, versión revisada	Carvajal, Hribernik, Duarte, Sanz-Rubiales y Centeno, 2013	Versión española del ESAS-r llevada a cabo por los autores del estudio. Contaron con profesionales bilingües con experiencia en cuidados paliativos y en el uso del ESAS, que elaboraron la propuesta de traducción del ESAS-r, basándose en la versión española del ESAS y la versión original del ESASr. Se utilizó el método de traducción inversa por parte de profesionales de la traducción bilingües. Se realizó una prueba piloto y un grupo de especialistas tuvo en cuenta los comentarios de los y las pacientes para crear la versión española final del ESAS-r para validar.  Aportan el cuestionario validado en el propio artículo.	66 pacientes con cáncer avanzado (al menos 30 de cada departamento de oncología de los hospitales terciarios participantes: Hospital de la Universidad de Navarra, España, e Instituto Nacional del Cáncer, Guatemala). 70 % mujeres. Edad media = 54 años. Rango de edad entre 18 y 84 años.
Escalas Visuales Analógicas (EVA) de ansiedad y depresión	Martínez, Porta, Espinosa y Gómez, 2006	No consta.	53 pacientes diagnosticados de cáncer avanzado o terminal que se atienden por primera vez en la consulta externa (CEXT) del Servicio de Cuidados Paliativos (SCP) del Instituto Catalán de Oncología (ICO). 66 % hombres. Edad media = 63 años (dt = 12,59).
Hospital Anxiety and Depression Scale (HADS; Zigmond y Snaith, 1983)  Escala Hospitalaria de Ansiedad y Depresión	López-Roig, Terol, Pastor, Neipp, Massutí, Rodríguez-Marín, Leyda, Martín-Aragón y Sitges, 2000	Versión hispana remitida por el equipo autor original.  Aportan el cuestionario validado en el propio artículo.	484 pacientes oncológicos. 69 % mujeres. Mayores de 20 años. Edad media = 55,7 años (dt = 13,3).  Ámbito: consultas externas (28,7 %); hospital de día (27,1 %); planta hospitalaria (23,1 %); o consulta de psicología (21,1 %).
	Costa, Pérez, Salamero y Gil, 2009	Versión española de la HADS (no especifica autoría).	234 pacientes oncológicos ambulatorios, que acudían al hospital de día para recibir tratamiento de quimioterapia o asistían a las consultas en las unidades funcionales. 51,7 % mujeres. Edad media = 55,24 años (dt = 11,5). Rango de edad entre 19 y 80 años.
	Martínez, Durá, Andreu, Galdón, Murgui e Ibañez, 2012	Versión española de Caro e Ibañez (1992).	892 pacientes oncológicos, que visitaban la Fundación Instituto Valenciano de Oncología (FIVO) por cualquier razón (diagnóstico, tratamiento, seguimiento). 70 % mujeres. Edad media = 57 años (dt = 13,3). Rango de edad entre 24 y 93 años.

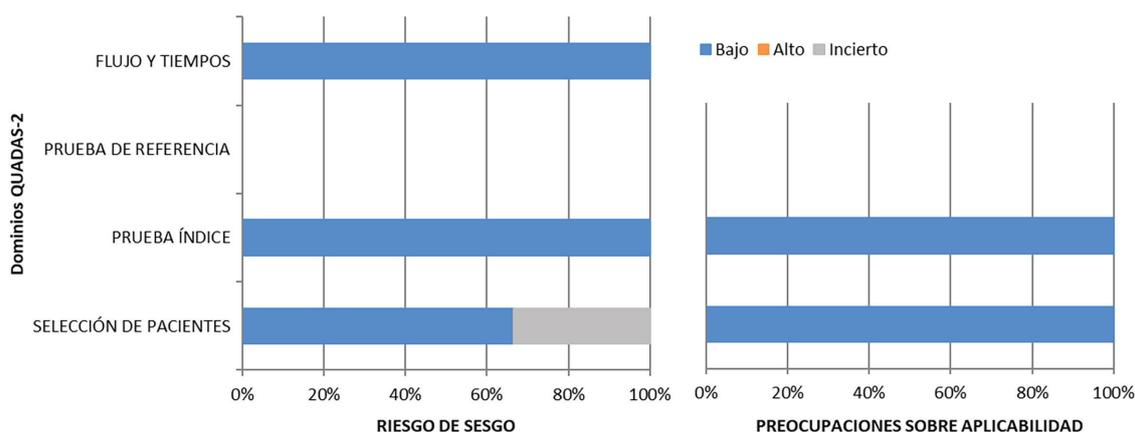
Instrumento original	Estudio de validación española en población oncológica	Versión española utilizada	Muestra validación española / Ámbito sanitario
Distress Thermometer (DT; Roth, Kornblith, Batel-Copel, Peabody, Scher y Holland, 1998)  Termómetro de distrés  Distress Thermometer (DT; Roth, Kornblith, Batel-Copel, Peabody, Scher y Holland, 1998)  Termómetro de distrés	Martínez, Galdón, Andreu e Ibáñez, 2013	Se entiende que han empleado la versión española vigente en ese momento según las <i>NCCN Guidelines</i> , pero no lo explicitan. En este estudio, se presentó la definición de malestar de la NCCN® antes de completar el DT, para brindar a los y las pacientes un marco de referencia.	962 pacientes con cáncer que acuden a la Fundación Instituto Valenciano de Oncología (FIVO) para recibir tratamiento en cualquiera de sus departamentos médicos. 68 % mujeres. Edad media = 57,55 años (dt = 13,34).
	Priede, Ruiz, Hoyuela, Herrán y González- Blanch, 2014	Versión española aportada por las <i>NCCN Guidelines</i> (2003).	135 pacientes con un primer diagnóstico reciente de cáncer confirmado por estudio de anatomía patológica que acuden al Hospital Universitario Marqués de Valdecilla (Santander, España). 65,19 % mujeres. Edad media = 52,84 años (dt = 10,15).
	Martínez, Andreu, Galdón e Ibáñez, 2015	Versión española del DT validada por Martínez, Galdón, Andreu e Ibáñez (2013).  Utilizan complementariamente el Impact Thermometer (IT) o Termómetro de impacto (Akizuki, Yamawaki, Akechi, Nakano, y Uchitomi, 2005) para explorar las posibilidades de mejora de la precisión diagnóstica con ambos instrumentos (tanto puntuaciones sumadas —DIT— como pareadas). El IT emplea una escala visual de 11 puntos (0 —Ausencia de interferencia— y 10 —Interferencia extrema—) para determinar la interferencia en la vida cotidiana del malestar emocional. No especifican qué versión española emplean.	385 pacientes oncológicos en edad adulta que visitan algún departamento médico de la FIVO. 55 % varones. Edad media = 58,7 años (dt = 13,2).
	Andreu, Galdón, Martínez y Martínez, 2015	Versión española del DT validada por Martínez, Galdón, Andreu e Ibáñez (2013).  Utilizan complementariamente la versión española del IT (Martínez, Andreu, Galdón e Ibáñez, 2015) para explorar las posibilidades de mejora de la precisión diagnóstica con ambos instrumentos (tanto puntuaciones sumadas —DIT— como pareadas). El IT emplea una escala visual de 11 puntos (0 —Ausencia de interferencia— y 10 —Interferencia extrema—) para determinar la interferencia en la vida cotidiana del malestar emocional.	81 pacientes oncológicos en edad adulta hospitalizados con neoplasias hematológicas en el Hospital Universitario la Fe de Valencia. 53 % varones. Edad media = 51 años (dt = 10,7).
	Andreu, Martínez y Galdón, 2019	Versión española del DT validada por Martínez, Galdón, Andreu e Ibáñez (2013) y Martínez, Andreu, Galdón e Ibáñez (2015).	236 pacientes en edad adulta supervivientes tras tratamiento de cáncer de inicio en la edad adulta, que acuden a la FIVO para citas médicas de seguimiento. 65 % mujeres. Edad media = 58 años (dt = 13,4).
Cuestionario de Detección de Malestar Emocional (DME; Maté, Mateo, Bayés, Bernaus, Casas, González-Barboteo, Bayés, Bernaus, Casas, López, Sirgo y Viel, 2009)	Limonero, Mateo, MatéMéndez, González-Barboteo, Bayés, Bernaus, Casas, López, Sirgo y Viel, 2012	Versión española original de Maté, Mateo, Bayés, Bernaus, Casas, González-Barboteo, Limonero, López-Postigo, Sirgo y Viel (2009).  Aportan el cuestionario validado en el propio artículo.	105 pacientes oncológico-paliativos hospitalizados en unidades de cuidados paliativos en cinco hospitales de Cataluña (Consorci Sanitari de Terrassa, Institut Català d'Oncologia, Corporació Sanitaria Parc Taulí de Sabadell, Hospital Universitari Sant Joan de Reus y la Fundació Hospital-Residència Sant Camil de Sant Pere de Ribes). 52 % mujeres. Edad media = 67,3 años (dt = 12,6).
Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida (ENP-E)	Mateo-Ortega, Limonero, Maté-Méndez, Beas, González-Barboteo, Barbero, Ela y Gómez-Batiste, 2019	Versión española original de Mateo-Ortega, Limonero, MatéMéndez, Beas, González-Barboteo, Barbero, Ela y Gómez-Batiste, 2019.	150 pacientes en edad adulta con diagnóstico de cáncer avanzado o terminal que ingresan en alguna de las unidades de cuidados paliativos de los 19 hospitales participantes: 21 pacientes formaron parte del estudio piloto; y 129 pacientes formaron parte de la fase de estudio. La información de un paciente fue descartada por estar incompleta. 51 % mujeres. Edad media en hombres = 72 años (rango: 33 - 93); edad media en mujeres = 73 años (rango: 43 - 93).

## Calidad de los estudios

La calidad de la evidencia de los estudios se valoró aplicando la herramienta QUADAS-2 para estudios de eficacia diagnóstica (89). En el anexo D puede consultarse la adaptación de QUADAS2 que se empleó, así como una tabla resumen de los resultados de la evaluación de la calidad de los estudios incluidos. En las siguientes figuras, se muestran los resultados resumidos en formato gráfico y agrupados según la herramienta de cribado analizada, y se indica para cada dominio tanto el porcentaje de estudios que tienen riesgo de sesgo bajo, alto o incierto, como el porcentaje de estudios con una preocupación baja, alta o incierta en cuanto a su aplicabilidad.

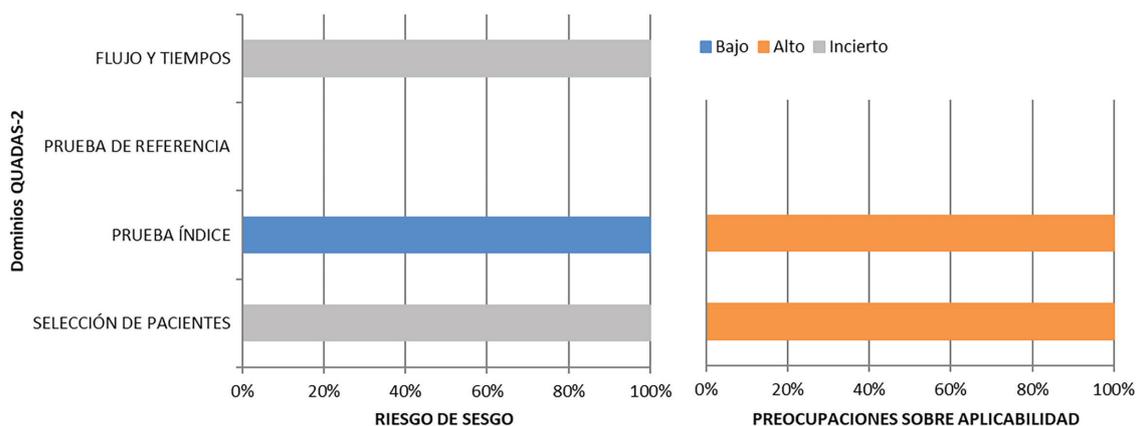
En general, los tres estudios de validación del BSI-18 (90, 92, 93) presentan un riesgo de sesgo bajo. Con respecto al dominio selección de pacientes, se estima un riesgo incierto en el caso del estudio de Galdón *et al.* (90), dado que no se aporta información sobre si la muestra incluida fue aleatoria o consecutiva. No se ha valorado el dominio prueba de referencia ni en cuanto al riesgo de sesgo ni a la aplicabilidad, puesto que estos estudios, tal y como se ha explicitado anteriormente, no cuentan con prueba de referencia. La aplicabilidad de los resultados no genera preocupaciones (figura 9).

**Figura 9. Evaluación de la calidad de los tres estudios sobre el BSI-18, según la herramienta QUADAS-2**



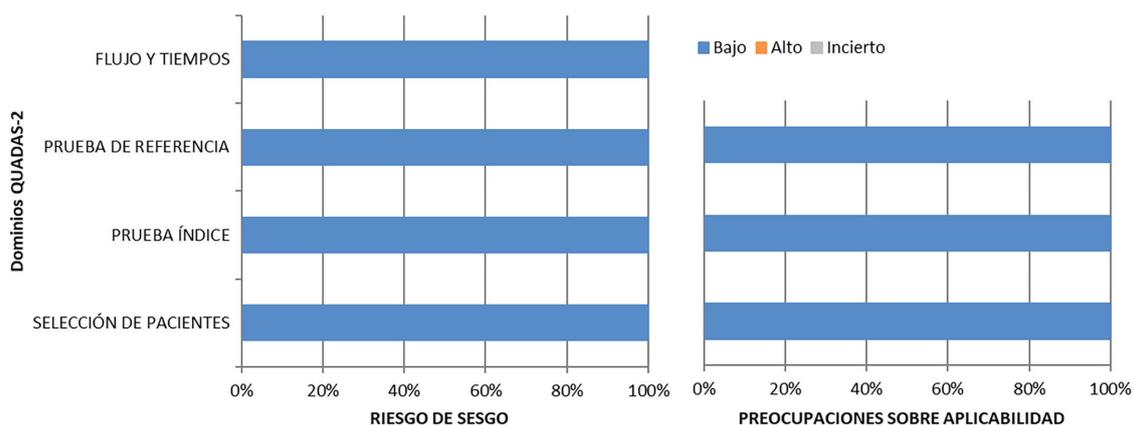
En el caso del estudio de Carvajal *et al.* sobre el ESAS-r (100), la selección de pacientes presenta un riesgo de sesgo incierto, ya que no se aporta información sobre si la muestra incluida fue aleatoria o consecutiva, ni qué cantidad de pacientes provienen de un centro u otro. Con respecto al flujo de pacientes, el riesgo de sesgo es incierto, dado que no se informa de pacientes que hayan declinado participar en el estudio ni de otras posibles exclusiones que hayan tenido lugar a lo largo del diagrama flujo. La preocupación sobre la aplicabilidad de los resultados es alta en lo que concierne a la selección de pacientes, puesto que la muestra es pequeña ( $n = 66$ ) y se desconoce la cantidad exacta que procede de España. La aplicabilidad de los resultados con respecto a la prueba índice también genera preocupaciones, ya que no se explicita el punto de corte de cada uno de los ítems o el total de la escala. No se valoró el dominio de prueba de referencia porque el estudio no cuenta con ello (figura 10).

**Figura 10. Evaluación de la calidad del estudio sobre el ESAS-r, según la herramienta QUADAS-2**



Por su parte, el estudio de Martínez *et al.* (101) sobre las Escalas Visuales Analógicas de ansiedad y depresión no genera preocupaciones respecto al riesgo de sesgo ni la aplicabilidad de los resultados (figura 11).

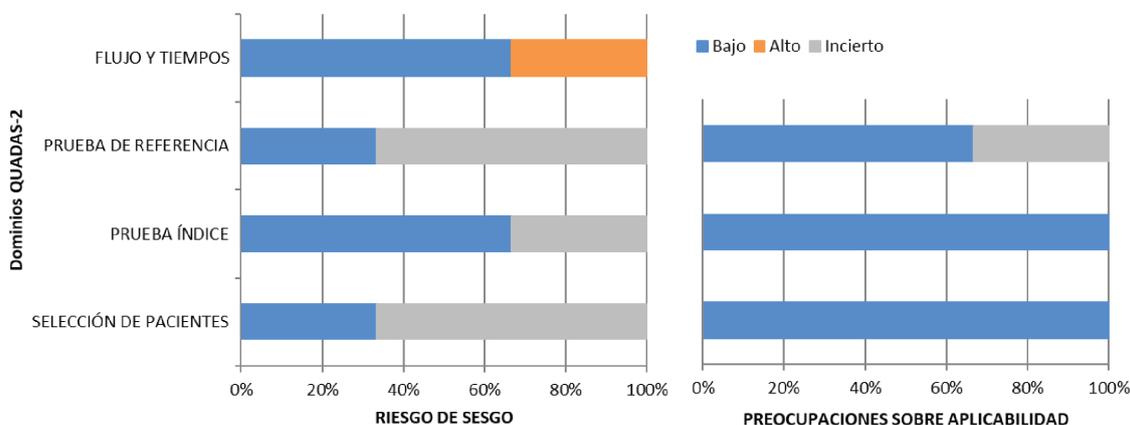
**Figura 11. Evaluación de la calidad del estudio sobre las EVA de ansiedad y depresión, según la herramienta QUADAS-2**



En el caso de los tres estudios sobre la Escala Hospitalaria de Ansiedad y Depresión (94-96), la selección de pacientes genera un riesgo de sesgo en dos de ellos (94, 95), pues no se menciona si la muestra es consecutiva o aleatoria. Además, en el caso del estudio de López-Roig *et al.* (94), se reclutaron pacientes en las consultas de psicología (21 % de la muestra), lo que pudiera dar lugar a una sobreestimación de la precisión diagnóstica de la prueba. La prueba índice no genera riesgo de sesgo en dos de los estudios, pero en el de López-Roig *et al.* (94) el riesgo resulta incierto, ya que se desconoce cuántas personas la recibieron de manera autoadministrada o heteroadministrada, si fue realizada antes o después de la prueba de referencia y si fue interpretada con o sin conocimientos de los resultados de la prueba de referencia. En lo que concierne a las pruebas de referencia, López-Roig *et al.* (94) no explicitan si se realizaron antes o después de la prueba índice ni si se aplicaron en formato autoadministrado o heteroadministrado, lo que hace estimar un riesgo incierto en este dominio. En el caso del estudio de Costa *et al.* (95), la prueba de referencia se realizó antes de la prueba índice y se desconoce si los resultados de la prueba

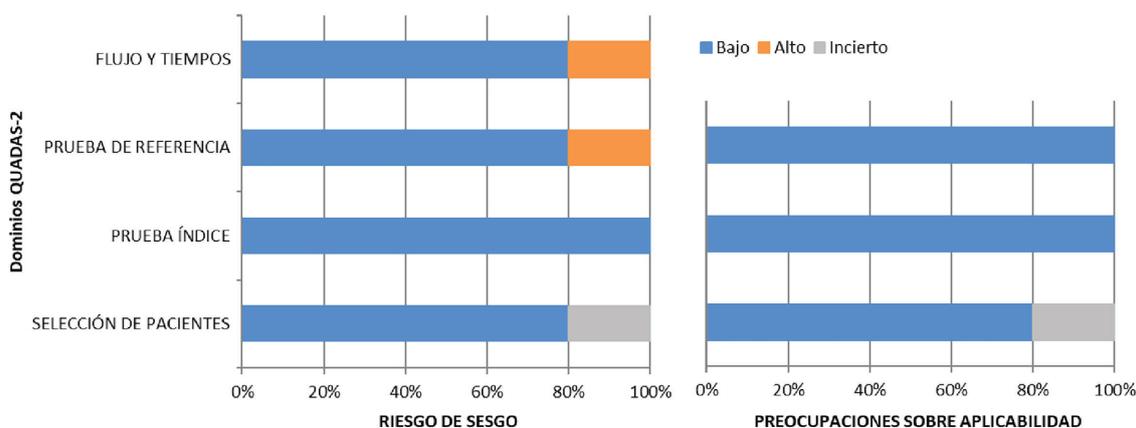
de referencia fueron interpretados con o sin conocimiento previo de los resultados de la prueba índice, por lo que también se estima un riesgo de sesgo incierto en este apartado. Con relación al flujo y tiempos, en el estudio de López-Roig *et al.* (94) no todas las personas pacientes reciben la misma prueba de referencia, y se desconocen las características de aquellas que no las recibieron. Además, tampoco se explicita cuál es el intervalo de tiempo que pasa entre la aplicación de la prueba índice y las pruebas de referencias, por lo que el riesgo de sesgo estimado es alto. No existen preocupaciones sobre la aplicabilidad, a excepción de la estimación de un riesgo incierto sobre la prueba de referencia del estudio de López-Roig *et al.* (94), dado que no aportan datos de estudios previos que avalen la fiabilidad, la validez y la precisión diagnóstica de las pruebas de referencia (STAI-S; SDS) (figura 12).

**Figura 12. Evaluación de la calidad de los tres estudios sobre la HADS, según la herramienta QUADAS-2**



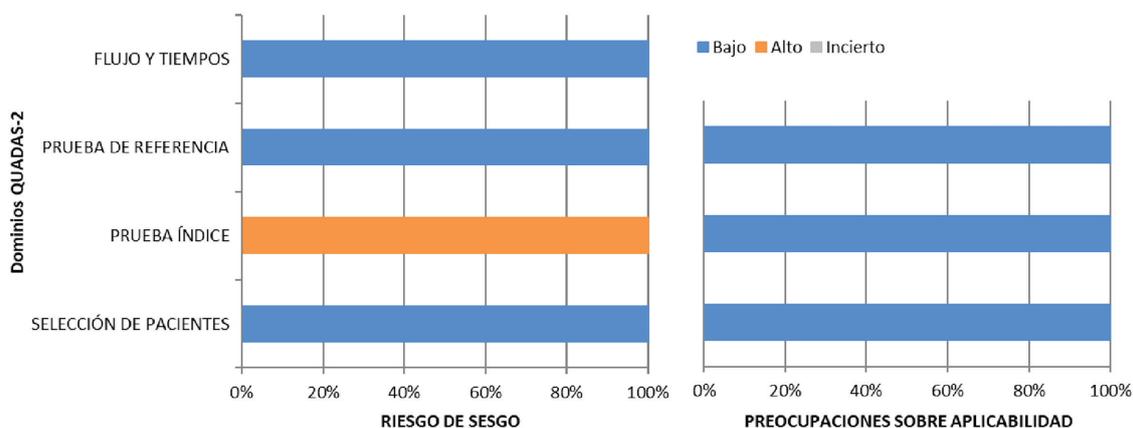
Con respecto al Termómetro de estrés, cuatro de los cinco estudios de validación seleccionados (35, 97-99) no generan preocupaciones en cuanto al riesgo de sesgo ni a la aplicabilidad. En el estudio de Priede *et al.* (91), la selección de pacientes genera preocupaciones respecto al riesgo de sesgo, pues no se aclara si se trata de una muestra consecutiva o aleatoria de pacientes, y se estima un riesgo de sesgo alto respecto al flujo y tiempos, ya que los y las pacientes seleccionados que finalmente no participan difieren significativamente de las que participan en nivel educativo y antecedentes psicopatológicos; además, el intervalo de tiempo que transcurre entre la prueba índice (DT) y la de referencia (HADS) se desconoce. Esto también genera incertezas sobre la aplicabilidad de los resultados en el dominio de selección de pacientes. Con relación a la prueba de referencia empleada, existe un riesgo de sesgo alto en dicho estudio (91), dado que se autoadministra en el domicilio de los y las pacientes, por lo que no se puede asegurar que haya sido la persona en tratamiento de manera individual quien ha respondido, además de que el ámbito de aplicación habitual de la HADS es el entorno sanitario (ambulatorio u hospitalario) y no el domiciliario (figura 13).

**Figura 13. Evaluación de la calidad de los cinco estudios sobre el DT, según la herramienta QUADAS-2**



El estudio sobre el Cuestionario de Detección de Malestar Emocional (102) no genera preocupaciones respecto al riesgo de sesgo, salvo con respecto a la prueba índice, pues no queda claro por qué se selecciona el punto de corte mayor o igual a 9, cuando al analizar el área bajo la curva ROC, los puntos óptimos indicados son 9,25 o 9,75 —dependiendo de la prueba de referencia empleada—. Además, no se aportan los valores de sensibilidad y especificidad de la escala en base a ese nuevo punto seleccionado, por lo que se desconoce su precisión diagnóstica concreta. No genera preocupaciones sobre la aplicabilidad de los resultados (figura 14).

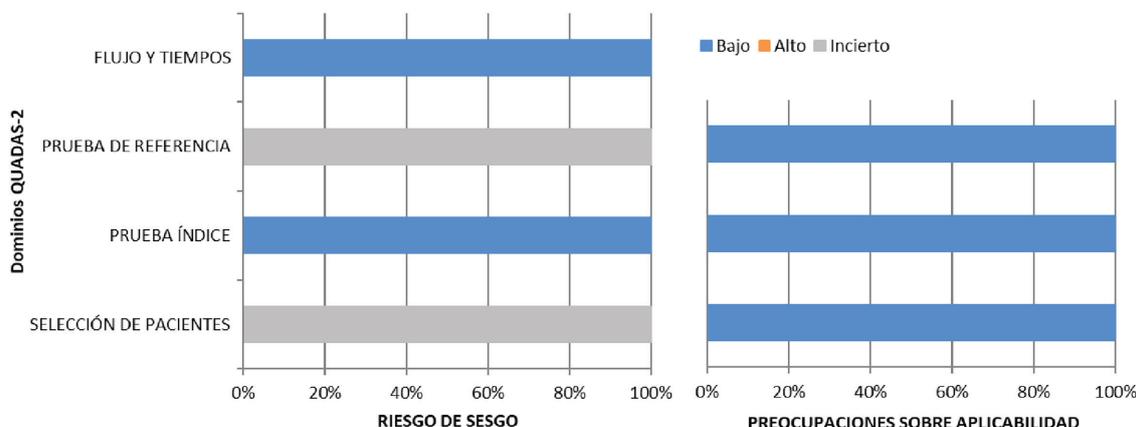
**Figura 14. Evaluación de la calidad del estudio sobre el DME, según la herramienta QUADAS-2**



Por su parte, en el estudio de validación del Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de Vida (103), no se explicita si la muestra incluida en el estudio es consecutiva o aleatoria, por lo que la selección de pacientes genera un riesgo de sesgo incierto. En relación con la prueba de referencia, está formada por una combinación de 4 instrumentos, aunque no se aclara la puntuación de corte de uno de ellos. Además, tampoco se menciona si las pruebas de referencia se realizaron antes o después de la prueba índice y se desconoce si se aplicaron en formato autoadministrado o heteroadministrado, por lo que el riesgo de sesgo en este dominio se valora como incierto.

El riesgo de sesgo estimado en el resto de los dominios es bajo, y no genera preocupaciones sobre la aplicabilidad de los resultados (figura 15).

**Figura 15. Evaluación de la calidad del estudio sobre el ENP-E, según la herramienta QUADAS-2**



## 4.2. Efectividad clínica de la tecnología

Dada la alta heterogeneidad de los índices y los resultados aportados por unos y otros estudios, se han diseñado unas tablas resumen que permiten visualizar de forma comparativa las propiedades psicométricas calculadas con más frecuencia en los estudios de validación, siguiendo lo que plantea el protocolo. Cabe mencionar que los resultados detallados de cada estudio se presentan en las tablas de evidencia (anexo B).

### 4.2.1. Propiedades psicométricas de las herramientas

En la primera tabla se ha recogido esta información de cada herramienta: el nombre del instrumento original junto con los autores y las autoras originales y la fecha de publicación; los estudios de validación española en población oncológica que se han seleccionado; los datos sobre la consistencia interna de la escala —como método más utilizado para medir la fiabilidad de los instrumentos—, y sobre los índices correspondientes a evidencias de validez que han sido empleados en al menos la mitad de los estudios —en concreto, los resultados sobre la validez de constructo, a través del análisis factorial, y las correlaciones que dan cuenta de la validez convergente—.

En el caso del BSI-18, los resultados de los tres estudios muestran que el cuestionario cuenta una consistencia interna aceptable (90, 92, 93) en la puntuación total, siendo esta inferior en cada una de las subescalas. Los tres estudios también confirman una solución jerárquica con tres factores de primer orden (subescalas de ansiedad, depresión y somatización), que había sido propuesta por el autor original (104), y un factor de segundo orden como medida de distrés emocional (GSI). Dos de los tres estudios exploran la validez convergente del inventario a través del cálculo de los coeficientes de correlación de la puntuación total con las cinco categorías de la Lista de Problemas y con la Escala Multidimensional de Apoyo Social Percibido. Todos estos resultados se pormenorizan en la tabla 10.

**Tabla 10. Consistencia interna y evidencias de validez del BSI-18 en los estudios españoles de validación en población oncológica**

Instrumento original	Estudio de validación española en población oncológica	Consistencia interna <sup>1</sup>	Validez de constructo: Análisis factorial e índices de bondad de ajuste		Validez convergente <sup>4</sup>
			AFE <sup>2</sup> Solución factorial; Índice KMO; Prueba de esfericidad de Bartlett	AFC <sup>3</sup> Solución factorial; $\chi^2/g$ ; RMSEA (IC 90 %); NNFI o TLI; CFI	
Brief Symptom Inventory-18 (BSI-18; Derogatis, 2000)  Inventario Breve de Síntomas 18 (BSI-18)	Galdón, Durá, Andreu, Ferrando, Murgui, Pérez e Ibañez, 2008	$\alpha$ GSI = 0,82 $\alpha$ Somatización = 0,62 $\alpha$ Depresión = 0,73 $\alpha$ Ansiedad = 0,72	-	Solución jerárquica con tres factores de primer orden (somatización, depresión y ansiedad) y un factor de segundo orden (GSI o puntuación total).  $\chi^2/g$ : 1,16; RMSEA (IC 90 %) = 0,03 (0,00 - 0,05); NNFI = 0,93	-
	Martínez, Conchado, Andreu y Galdón, 2019	$\alpha$ GSI = 0,91 $\alpha$ Somatización = 0,72 $\alpha$ Depresión = 0,83 $\alpha$ Ansiedad = 0,84	-	Solución jerárquica con tres factores de primer orden (somatización, depresión y ansiedad) y un factor de segundo orden (GSI o puntuación total).  $\chi^2/g$ : 2,82; RMSEA (IC 90 %) = 0,039 (0,035 - 0,044); NNFI = 0,915; CFI = 0,927	Correlaciones positivas estadísticamente significativas entre el BSI-18 (GSI) y las cinco categorías de la Lista de Problemas (problemas prácticos = 0,29; problemas familiares = 0,38; problemas emocionales = 0,66; problemas espirituales o religiosos = 0,13; y problemas físicos = 0,57). También correlaciones positivas estadísticamente significativas entre el BSI-18 (subescalas) y las cinco categorías de la Lista de Problemas.
	Calderon, Ferrando, Lorenzo-Seva, Hernández, Oporto-Alonso y Jiménez-Fonseca, 2020	$\alpha$ GSI = 0,91 $\alpha$ Somatización = 0,72 $\alpha$ Depresión = 0,83 $\alpha$ Ansiedad = 0,89	Solución jerárquica de segundo orden, con tres factores primarios fuertemente relacionados (somatización, depresión y ansiedad) y un factor general de segundo orden (GSI).  Informan de que el índice KMO y la prueba de esfericidad de Bartlett son adecuados, pero no aportan los datos concretos.	Solución jerárquica de segundo orden con tres factores primarios (subescalas de somatización, depresión y ansiedad).  $\chi^2/g$ : 2,84; RMSEA (IC 90 %) = 0,046 (0,041 - 0,051); CFI = 0,978	Coefficiente de correlación negativo estandarizado y estadísticamente significativo entre el BSI-18 (GSI) y la Escala Multidimensional de Apoyo Social Percibido (MSPSS) = - 0,21. También correlaciones negativas estadísticamente significativas entre el BSI-18 (subescalas) y el MSPSS (subcategorías de familia, amistades y personas importantes).

<sup>1</sup>  $\alpha < 0,7$  muestra que la consistencia interna de la escala es baja;  $0,7 \leq \alpha \leq 0,9$  se consideran valores aceptables;  $\alpha > 0,9$  se considera que hay redundancia o duplicación de ítems (116).

<sup>2</sup> La prueba de esfericidad de Bartlett pone a prueba la hipótesis nula que propone que la matriz de correlaciones entre los ítems de la escala es una matriz identidad (es decir, que todas las correlaciones son 0), pues no tendría sentido buscar factores que expliquen la varianza compartida entre los ítems si estos no comparten varianza. Esto es, si la prueba resulta significativa ( $p < 0,05$ ), se puede aplicar el análisis factorial; si no lo es, no se debe aplicar el análisis factorial. Índice KMO: indica el grado de intercorrelación de las variables. Se siguen los criterios de valoración propuestos por Kaiser & Rice (117):  $KMO \geq 0,9$  se considera excelente;  $0,9 > KMO \geq 0,8$  sería bueno;  $0,8 > KMO \geq 0,7$  es aceptable;  $0,7 > KMO \geq 0,6$  se considera mediocre;  $0,6 > KMO \geq 0,5$  sería bajo;  $KMO < 0,50$ , inaceptable.

<sup>3</sup> Índices de bondad de ajuste absolutos:  $\chi^2/g$ : según Wheaton et al. (118), un cociente de 5 indica un ajuste razonable, mientras que Brooke et al. (119) señalan que los valores inferiores o en torno a 2 indican un ajuste excelente. RMSEA: los valores inferiores o cercanos a 0,06 se consideran adecuados (120).

Índices de bondad de ajuste comparativos: para los estadísticos NNFI (o TLI) y CFI, los valores por encima del 0,90 suponen un ajuste aceptable (121, 122), y los superiores o iguales a 0,95 indican un ajuste óptimo (120).

<sup>4</sup> Interpretación de los coeficientes de correlación siguiendo a Hinkle, Wiersma & Jurs (123): entre 0 y 0,3 = correlación inexistente o escasa; entre 0,3 y 0,5 = correlación débil; entre 0,5 y 0,7 = correlación moderada; entre 0,7 y 0,9 = correlación fuerte; entre 0,9 y 1 = correlación muy fuerte.

**Abreviaturas:** Fundación Instituto Valenciano de Oncología (FIVO); Desviación típica (dt); Índice Global de Severidad (GSI); Análisis factorial exploratorio (AFE); Análisis factorial confirmatorio (AFC); Índice Kaiser-Mayer-Olkin (KMO); Grados de libertad (gl); Root Mean Square Error of Approximation (RMSEA); Intervalo de confianza (IC); Non-Normed Fit Index (NNFI) o Tucker-Lewis Index (TLI); Comparative Fit Index (CFI); Alfa de Cronbach ( $\alpha$ ).

**Fuente:** Elaboración propia.

En el caso del ESAS-r, en el estudio de validación de Carvajal *et al.* (100) se incluyó el cálculo de la consistencia interna, siendo esta buena ( $\alpha = 0,86$ ), pero no se llevó a cabo ningún tipo de análisis factorial ni se analizó la validez convergente de la escala.

Con referencia a las EVA de ansiedad y depresión, no se calculó ni la consistencia interna ni la validez de constructo en el estudio de validación (101), entendiéndolas como medidas complementarias pero no pertenecientes a una misma escala. Tampoco se calcularon otros indicadores de fiabilidad (como, por ejemplo, la estabilidad temporal). Con relación a la validez convergente, la correlación entre el Distrés Emocional Global (DEG) o la puntuación obtenida de la suma de las puntuaciones de las EVA y la puntuación total de la HADS (0,77) fue fuerte, positiva y estadísticamente significativa. Las correlaciones calculadas entre la EVA de ansiedad y la subescala de la HADS de ansiedad (0,663) y entre la EVA de depresión y la subescala de la HADS de depresión (0,576) resultaron positivas, moderadas y estadísticamente significativas.

En lo que concierne a la HADS, los tres estudios de validación (94-96) informaron de una consistencia interna adecuada, con valores superiores a 0,80 en todos los casos. En dos de los estudios (94, 95) realizaron un Análisis Factorial Exploratorio (AFE) cuyo resultado fue una solución bifactorial, mientras que en el estudio de Martínez *et al.* (96) se llevó a cabo un Análisis Factorial Confirmatorio (AFC) donde se comparó la solución bifactorial propuesta por los autores originales con una solución jerárquica con dos factores de primer orden (subescalas de depresión y ansiedad), y un factor de segundo orden (puntuación total). Aunque ambos modelos se ajustan correctamente a los datos, los índices de ajuste muestran que la solución jerárquica con dos factores de primer orden y uno de segundo orden es la estructura más apropiada para la validación en población oncológica española (96). Uno de los estudios (94) también analiza la validez convergente de la escala. Estos resultados se detallan en la tabla 11.

**Tabla 11. Consistencia interna y evidencias de validez de la HADS en los estudios españoles de validación en población oncológica**

Instrumento original	Estudio de validación española en población oncológica	Consistencia interna <sup>1</sup>	Validez de constructo: Análisis factorial e índices de bondad de ajuste		Validez convergente <sup>4</sup>
			AFE <sup>2</sup> Solución factorial; Índice KMO; Prueba de esfericidad de Bartlett	AFC <sup>3</sup> Solución factorial; $\chi^2/gl$ ; RMSEA (IC 90 %); NNFI o TLI; CFI	
Hospital Anxiety and Depression Scale (HADS; Zigmond y Snaith, 1983)  Escala Hospitalaria de Ansiedad y Depresión	López-Roig, Terol, Pastor, Neipp, Massutí, Rodríguez-Marín, Leyda, Martín-Aragón y Sitges, 2000	$\alpha$ HADS-A = 0,85 $\alpha$ HADS-D = 0,87	Solución bifactorial (depresión y ansiedad) Índice KMO = 0,935 Prueba de esfericidad de Bartlett: $\chi^2 = 2955,6$ ( $p < 0,001$ )	-	Coefficiente de correlación de Spearman positivo y estadísticamente significativo entre la HADS-A y el Inventario de Ansiedad Estado-Rasgo (subescala de ansiedad-estado: STAI-S) = 0,71; y entre la HADS-D y la Escala de depresión de Zung (SDS) = 0,76
	Costa, Pérez, Salamero y Gil, 2009	$\alpha$ HADS-A = 0,82 $\alpha$ HADS-D = 0,84	Solución bifactorial (depresión y ansiedad) Índice KMO = 0,91 Prueba de esfericidad de Bartlett: $\chi^2 = 1242,86$ ( $p < 0,001$ )	-	-
	Martínez, Durá, Andreu, Galdón, Murgui e Ibañez, 2012	$\alpha$ HADS-total = 0,89 $\alpha$ HADS-A = 0,83 $\alpha$ HADS-D = 0,83	-	Solución jerárquica con dos factores de primer orden (subescalas de depresión y ansiedad), y un factor de segundo orden (puntuación total) $\chi^2/gl = 2,36$ ; RMSEA (IC 90 %) = 0,04 (0,03 - 0,05); NNFI = 0,96	-

<sup>1</sup>  $\alpha < 0,7$  muestra que la consistencia interna de la escala es baja;  $0,7 \leq \alpha \leq 0,9$  se consideran valores aceptables;  $\alpha > 0,9$  se considera que hay redundancia o duplicación de ítems (116).

<sup>2</sup> La prueba de esfericidad de Bartlett pone a prueba la hipótesis nula que propone que la matriz de correlaciones entre los ítems de la escala es una matriz identidad (es decir, que todas las correlaciones son 0), pues no tendría sentido buscar factores que expliquen la varianza compartida entre los ítems si estos no comparten varianza. Esto es, si la prueba resulta significativa ( $p < 0,05$ ), se puede aplicar el análisis factorial; si no lo es, no se debe aplicar el análisis factorial. Índice KMO: indica el grado de intercorrelación de las variables. Se siguen los criterios de valoración propuestos por Kaiser & Rice (117):  $KMO \geq 0,9$  se considera excelente;  $0,9 > KMO \geq 0,8$  sería bueno;  $0,8 > KMO \geq 0,7$  es aceptable;  $0,7 > KMO \geq 0,6$  se considera mediocre;  $0,6 > KMO \geq 0,5$  sería bajo;  $KMO < 0,50$ , inaceptable.

<sup>3</sup> Índices de bondad de ajuste absolutos:  $\chi^2/gl$ : según Wheaton et al. (118), un cociente de 5 indica un ajuste razonable, mientras que Brooke et al. (119) señalan que los valores inferiores o en torno a 2 indican un ajuste excelente. RMSEA: los valores inferiores o cercanos a 0,06 se consideran adecuados (120).

Índices de bondad de ajuste comparativos: para los estadísticos NNFI (o TLI) y CFI: los valores por encima del 0,90 suponen un ajuste aceptable (121, 122) y los superiores o iguales a 0,95 indican un ajuste óptimo (120).

<sup>4</sup> Interpretación de los coeficientes de correlación siguiendo a Hinkle, Wiersma & Jurs (123): entre 0 y 0,3 = correlación inexistente o escasa; entre 0,3 y 0,5 = correlación débil; entre 0,5 y 0,7 = correlación moderada; entre 0,7 y 0,9 = correlación fuerte; entre 0,9 y 1 = correlación muy fuerte.

**Abreviaturas:** Desviación típica (dt); Subescala de ansiedad de la HADS (HADS-A); Subescala de depresión de la HADS (HADS-D); Análisis factorial exploratorio (AFE); Análisis factorial confirmatorio (AFC); Índice Kaiser-Mayer-Olkin (KMO); Grados de libertad (gl); Root Mean Square Error of Approximation (RMSEA); Intervalo de confianza (IC); Non-Normed Fit Index (NNFI) o Tucker-Lewis Index (TLI); Comparative Fit Index (CFI); Alfa de Cronbach ( $\alpha$ ).

**Fuente:** Elaboración propia.

Con relación al DT, puesto que el instrumento consta de un único ítem, no se calculó ni su consistencia interna ni su estructura factorial. Tampoco se calcularon otros indicadores de fiabilidad (como, por ejemplo, la estabilidad temporal). La validez convergente del test se evaluó en todos los estudios de validación, informando de correlaciones positivas moderadas con medidas como el BSI-18, los problemas emocionales recogidos en la Lista de Problemas, o la HADS. Los resultados pormenorizados se muestran en la tabla 12.

**Tabla 12. Evidencias de validez del DT en los estudios españoles de validación en población oncológica**

Instrumento original	Estudio de validación española en población oncológica	Consistencia interna	Validez de constructo: Análisis factorial e índices de bondad de ajuste		Validez convergente <sup>1</sup>
			AFE Solución factorial; Índice KMO; Prueba de esfericidad de Bartlett	AFC Solución factorial; $\chi^2/gl$ ; RMSEA (IC 90 %); NNFI o TLI; CFI	
Distress Thermometer (DT; Roth, Kornblith, Batel-Copel, Peabody, Scher y Holland, 1998)	Martínez, Galdón, Andreu e Ibáñez, 2013	-	-	-	Correlaciones positivas estadísticamente significativas entre la puntuación del DT y el BSI-18 (GSI) = 0,65; así como con las subescalas: ansiedad (0,61), depresión (0,58) y somatización (0,41). También correlaciones positivas estadísticamente significativas entre el DT y los problemas emocionales recogidos en la Lista de Problemas: tristeza (0,53); preocupación (0,47); miedos (0,42); depresión (0,41); nerviosismo (0,41), y pérdida de interés (0,35).
	Priede, Ruiz, Hoyuela, Herrán y González Blanch, 2014	-	-	-	Correlación positiva estadísticamente significativa entre el DT y la HADS (0,64). También correlación positiva y estadísticamente significativa entre el DT y la HADS-A (0,66), y con la HADS-D (0,51).
	Martínez, Andreu, Galdón e Ibáñez, 2015	-	-	-	Coefficiente de correlación de Spearman positivo estadísticamente significativo entre el BSI-18 y el DT (0,62), y con la combinación DIT (0,63).
Termómetro de distrés	Andreu, Galdón, Martínez y Martínez, 2015	-	-	-	Coefficiente de correlación de Spearman positivo y estadísticamente significativo entre el BSI-18 y el DT (0,58), y con la combinación DIT (0,53). También correlaciones positivas y estadísticamente significativas entre el DT y las tres subescalas del BSI-18: ansiedad (0,57), depresión (0,54) y somatización (0,38), así como entre la combinación DIT y las tres subescalas del BSI-18: ansiedad (0,46), depresión (0,51) y somatización (0,38).
	Andreu, Martínez y Galdón, 2019	-	-	-	Coefficiente de correlación de Spearman positivo y estadísticamente significativo entre el DT y el BSI-18 (0,67), y con las tres subescalas: somatización (0,42), depresión (0,56) y ansiedad (0,62).

<sup>1</sup> Interpretación de los coeficientes de correlación siguiendo a Hinkle, Wiersma & Jurs (123): entre 0 y 0,3 = correlación inexistente o escasa; entre 0,3 y 0,5 = correlación débil; entre 0,5 y 0,7 = correlación moderada; entre 0,7 y 0,9 = correlación fuerte; entre 0,9 y 1 = correlación muy fuerte.

**Abreviaturas:** Termómetro de distrés (DT); Subescala de ansiedad de la HADS (HADS-A); Subescala de depresión de la HADS (HADS-D); Análisis factorial exploratorio (AFE); Análisis factorial confirmatorio (AFC); Índice Kaiser-Mayer-Olkin (KMO); Grados de libertad (gl); Root Mean Square Error of Approximation (RMSEA); Intervalo de confianza (IC); Non-Normed Fit Index (NNFI) o Tucker-Lewis Index (TLI); Comparative Fit Index (CFI); Alfa de Cronbach ( $\alpha$ ).

**Fuente:** Elaboración propia.

Con respecto al DME, en el estudio de validación de Limonero *et al.* (102) se obtuvo un valor de  $\alpha = 0,685$ , lo que indica que la consistencia interna de la escala es baja. A lo sumo, este valor podría considerarse cercano al valor mínimo aceptable (0,7), sobre todo teniendo en cuenta que está formada únicamente por dos ítems. La validez convergente de la escala se estudió a través del cálculo de los coeficientes de correlación intraclase con el DT (0,235), la HADS (0,513) y sus subescalas de ansiedad (0,614) y depresión (0,499), así como con una valoración psicológica clínica (0,438). Todos los coeficientes resultaron estadísticamente significativos.

En relación con el ENP-E, en el estudio de validación (103) se analizó la consistencia interna del total de la escala (sección A), siendo esta aceptable. También se realizó un AFE de dicha sección (A), cuyo resultado fue una solución cuatrifactorial, respaldada por índices de ajuste aceptables (103). No obstante, los bajos índices de confiabilidad de los factores de apoyo social y espiritualidad desaconsejan el empleo de las dimensiones por separado. No se calculó el alfa de Cronbach del factor información por estar formado por un único ítem. En cuanto a la validez convergente de la escala, se informó de coeficientes de correlación positiva moderada y estadísticamente significativa con la HADS y el DT.

**Tabla 13. Consistencia interna y evidencias de validez del ENP-E en los estudios españoles de validación en población oncológica**

Instrumento original	Estudio de validación española en población oncológica	Consistencia interna <sup>1</sup>	Validez de constructo: Análisis factorial e índices de bondad de ajuste		Validez convergente <sup>3</sup>
			AFE <sup>2</sup> Solución factorial; Índice KMO; Prueba de esfericidad de Bartlett	AFC Solución factorial; $\chi^2/gl$ ; RMSEA (IC 90 %); NNFI o TLI; CFI	
Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida (ENP-E)	Mateo-Ortega, Limonero, MatéMéndez, Beas, González-Barboteo, Barbero, Ela y GómezBatiste, 2019	$\alpha$ Total sección A = 0,75 $\alpha$ Bienestar emocional = 0,77 $\alpha$ Apoyo social = 0,64 $\alpha$ Espiritualidad = 0,48	Solución cuatrifactorial (bienestar emocional, apoyo social, espiritualidad e información); Índice KMO = 0,75; Prueba de esfericidad de Bartlett ( $p < 0,01$ ).	-	Índice de correlación intraclase (ICC) estadísticamente significativo entre el ENP-E y la HADS (0,67); con el ítem de calidad de vida del EORTC QLQ-C15-PAL (-0,0344); y con el DT (0,516). No resultó estadísticamente significativo con la percepción subjetiva del paso del tiempo (0,326).

<sup>1</sup>  $\alpha < 0,7$  muestra que la consistencia interna de la escala es baja;  $0,7 \leq \alpha \leq 0,9$  se consideran valores aceptables;  $\alpha > 0,9$  se considera que hay redundancia o duplicación de ítems (116).

<sup>2</sup> La prueba de esfericidad de Bartlett pone a prueba la hipótesis nula que propone que la matriz de correlaciones entre los ítems de la escala es una matriz identidad (es decir, que todas las correlaciones son 0), pues no tendría sentido buscar factores que expliquen la varianza compartida entre los ítems si estos no comparten varianza. Esto es, si la prueba resulta significativa ( $p < 0,05$ ), se puede aplicar el análisis factorial; si no lo es, no se debe aplicar el análisis factorial. Índice KMO: indica el grado de intercorrelación de las variables. Se siguen los criterios de valoración propuestos por Kaiser & Rice (117):  $KMO \geq 0,9$  se considera excelente;  $0,9 > KMO \geq 0,8$  sería bueno;  $0,8 > KMO \geq 0,7$  es aceptable;  $0,7 > KMO \geq 0,6$  se considera mediocre;  $0,6 > KMO \geq 0,5$  sería bajo;  $KMO < 0,50$ , inaceptable.

<sup>3</sup> Interpretación de los coeficientes de correlación siguiendo a Hinkle, Wiersma & Jurs (123): entre 0 y 0,3 = correlación inexistente o escasa; entre 0,3 y 0,5 = correlación débil; entre 0,5 y 0,7 = correlación moderada; entre 0,7 y 0,9 = correlación fuerte; entre 0,9 y 1 = correlación muy fuerte.

**Abreviaturas:** Análisis factorial exploratorio (AFE); Análisis factorial confirmatorio (AFC); Índice Kaiser-Mayer-Olkin (KMO); Grados de libertad (gl); Root Mean Square Error of Approximation (RMSEA); Intervalo de confianza (IC); Non-Normed Fit Index (NNFI) o Tucker-Lewis Index (TLI); Comparative Fit Index (CFI); Alfa de Cronbach ( $\alpha$ ).

**Fuente:** Elaboración propia.

En resumen, los estudios de validación del BSI-18 informan de una consistencia interna aceptable ( $\alpha$  entre 0,82 y 0,91), aportan evidencias de validez convergente y confirman una estructura jerárquica con tres factores de primer orden formada por las subescalas de ansiedad, depresión y somatización. Por su parte, el ESAS-r también cuenta con una consistencia interna aceptable ( $\alpha = 0,86$ ), pero no analiza ni la validez de constructo ni la validez convergente de la escala. Con referencia a la HADS, los valores aportados por los estudios de validación sobre su consistencia interna son aceptables ( $\alpha > 0,80$ ) y la estructura subyacente encontrada en dos de los estudios es bifactorial, formada por las subescalas de ansiedad y depresión. El otro estudio de validación encuentra una solución jerárquica con dos factores de primer orden (ansiedad y depresión), y un factor de segundo orden (puntuación total) como medida de estrés emocional. Asimismo, uno de los estudios presenta evidencias de validez convergente de cada subescala. Los estudios de validación del DT no calculan ni su consistencia interna ni su estructura factorial, dado que está formado por un solo ítem, y tampoco lo hace el estudio de validación de las EVA de ansiedad y depresión. Los estudios de validación del DT aportan evidencias de validez convergente a través de las correlaciones moderadas con las puntuaciones del BSI-18, la HADS y los síntomas emocionales de la Lista de Problemas, mientras que las EVA de ansiedad y depresión muestran una correlación fuerte con la HADS. En lo que concierne al DME, se calculó su consistencia interna, siendo esta baja ( $\alpha = 0,685$ ), y sus puntuaciones correlacionaron de manera moderada con el DT, la HADS y sus subescalas, y con una valoración psicológica clínica. En el caso del ENP-E, el estudio de validación informa de una consistencia interna aceptable ( $\alpha = 0,75$ ) y de una estructura cuatrifactorial subyacente, formada por las subescalas de bienestar emocional, apoyo social, espiritualidad e información. Sus puntuaciones correlacionaron de forma moderada con las de la HADS y el DT.

#### 4.2.2. Precisión diagnóstica de las herramientas

De manera complementaria, en otro grupo de tablas se ha recopilado la información más relevante con respecto a la precisión diagnóstica de cada escala, a través del análisis del área bajo la curva ROC. Este análisis es un método estadístico para pruebas diagnósticas que utilizan escalas continuas mediante el que se determina el punto de corte en el que se alcanza la sensibilidad y la especificidad más alta, se evalúa la capacidad discriminativa de la prueba diagnóstica y se compara la capacidad discriminativa de dos o más pruebas que expresan sus resultados como escalas continuas (124). En estas tablas, se recoge cuál fue la prueba de referencia empleada, el punto de corte establecido, la sensibilidad, la especificidad, el valor predictivo positivo (VPP), el valor predictivo negativo (VPN), la utilidad clínica positiva (CUI+), la utilidad clínica negativa (CUI-) y el valor global de la prueba. En este sentido, puesto que el objetivo es encontrar una herramienta de *screening* para el malestar emocional, los estudios han tratado de seleccionar aquellos puntos de corte que permiten alcanzar valores altos de porcentaje de pacientes correctamente clasificados (valor global de la prueba), pero haciendo hincapié en maximizar los valores de sensibilidad de la herramienta (125). Se incorporan a las tablas aquellos índices de los que informa cada estudio de validación.

En primer lugar, los estudios sobre el BSI-18 y el ESAS-r no emplearon ninguna prueba de referencia para medir su precisión diagnóstica. En el caso del BSI-18, el punto de corte se estableció siguiendo las puntuaciones estandarizadas propuestas por el autor original (104); en cuanto al ESAS-r, no se indicó el punto de corte.

**Tabla 14. Precisión diagnóstica del BSI-18, según los estudios de validación española en población oncológica**

Instrumento original	Estudio de validación española	Prueba de referencia	Punto de corte	Sensibilidad	Especificidad	AUC (IC 95 %) <sup>1</sup>	VPP	VPN	CUI+	CUI-	Valor global de la prueba
Brief Symptom Inventory-18 (BSI-18; Derogatis, 2000)	Galdón, Durá, Andreu, Ferrando, Murgui, Pérez e Ibañez, 2008	-	Las puntuaciones brutas del BSI-18 se convierten en puntuaciones T estandarizadas. Se considera que una persona presenta malestar emocional si su puntuación total es igual o superior a 63 en el GSI, o en al menos dos de las dimensiones, de acuerdo con el punto de corte recomendado por Derogatis (2000).	-	-	-	-	-	-	-	-
	Martínez, Conchado, Andreu y Galdón, 2019	-		-	-	-	-	-	-	-	-
Inventario Breve de Síntomas-18 (BSI-18)	Calderon, Ferrando, LorenzoSeva, Hernández, OportoAlonso y JiménezFonseca, 2020	-		-	-	-	-	-	-	-	-

<sup>1</sup> AUC ≥ 0,9 muestra una capacidad discriminativa del test excelente; 0,9 > AUC ≥ 0,8 se considera buena; 0,8 > AUC ≥ 0,7 es aceptable; 0,7 > AUC ≥ 0,6 se considera mediocre; 0,6 > AUC > 0,5 sería fallida (126).  
**Abreviaturas:** Índice Global de Severidad (GSI); Área bajo la curva (AUC); Intervalo de confianza (IC); Valor predictivo positivo (VPP); Valor predictivo negativo (VPN); Utilidad clínica positiva (CUI+); Utilidad clínica negativa (CUI-).  
**Fuente:** Elaboración propia.

**Tabla 15. Precisión diagnóstica del ESAS-r, según el estudio de validación española en población oncológica**

Instrumento original	Estudio de validación española	Prueba de referencia	Punto de corte	Sensibilidad	Especificidad	AUC (IC 95 %) <sup>1</sup>	VPP	VPN	CUI+	CUI-	Valor global de la prueba
Edmonton Symptom Assessment System-revised (ESAS-r; Watanabe, Nikolaichuk, Beaumont, Johnson, Myers y Strasser, 2011)  Cuestionario de evaluación de síntomas de Edmonton, versión revisada	Carvajal, Hribernik, Duarte, SanzRubiales y Centeno, 2013	-	No informan del punto de corte.  Las y los pacientes con mayores puntuaciones tienen peor estado funcional.	-	-	-	-	-	-	-	-

<sup>1</sup> AUC ≥ 0,9 muestra una capacidad discriminativa del test excelente; 0,9 > AUC ≥ 0,8 se considera buena; 0,8 > AUC ≥ 0,7 es aceptable; 0,7 > AUC ≥ 0,6 se considera mediocre; 0,6 > AUC > 0,5 sería fallida (126).  
**Abreviaturas:** Área bajo la curva (AUC); Intervalo de confianza (IC); Valor predictivo positivo (VPP); Valor predictivo negativo (VPN); Utilidad clínica positiva (CUI+); Utilidad clínica negativa (CUI-).  
**Fuente:** Elaboración propia.

Para evaluar la capacidad diagnóstica de las EVA de ansiedad y depresión, se empleó la HADS como prueba de referencia. El punto de corte establecido (como sumatorio de la puntuación de ambas escalas) fue 10 (101).

**Tabla 16. Precisión diagnóstica de las EVA de ansiedad y depresión, según el estudio de validación española en población oncológica**

Instrumento original	Estudio de validación española	Prueba de referencia	Punto de corte	Sensibilidad	Especificidad	AUC (IC 95 %) <sup>1</sup>	VPP	VPN	CUI+	CUI-	Valor global de la prueba
Escalas Visuales Analógicas (EVA) de ansiedad y depresión	Martínez, Porta, Espinosa y Gómez, 2006	Escala Hospitalaria de Ansiedad y Depresión (HADS). Compuesta por 14 ítems y dos subescalas (ansiedad y depresión). No se especifica qué versión emplea. Se selecciona el punto de corte 16 (como sumatorio de ambas subescalas).	≥ 10	0,80	0,939	0,870	-	-	-	-	-

<sup>1</sup> AUC ≥ 0,9 muestra una capacidad discriminativa del test excelente; 0,9 > AUC ≥ 0,8 se considera buena; 0,8 > AUC ≥ 0,7 es aceptable; 0,7 > AUC ≥ 0,6 se considera mediocre; 0,6 > AUC > 0,5 sería fallida (126).  
**Abreviaturas:** Área bajo la curva (AUC); Intervalo de confianza (IC); Valor predictivo positivo (VPP); Valor predictivo negativo (VPN); Utilidad clínica positiva (CUI+); Utilidad clínica negativa (CUI-).  
**Fuente:** Elaboración propia.

Con respecto a los estudios de validación de la HADS en población oncológica en España, se emplearon diferentes pruebas de referencia y se propusieron distintos puntos de corte. Para la puntuación total, que es la medida de malestar emocional propiamente dicha, Costa *et al.* estimaron un punto de corte ≥ 10 para discriminar el malestar emocional clínicamente significativo, mientras que Martínez *et al.* (96) indicaron que el punto de corte óptimo para detectar el malestar emocional general en los departamentos de oncología es ≥16.

Con relación a los estudios de validación del Termómetro de distrés, cuatro de los cinco estudios coinciden en señalar el punto de corte en ≥ 5 (35, 97-99), siendo los valores de sensibilidad cercanos o superiores al 90 %.

**Tabla 17. Precisión diagnóstica de la HADS, según los estudios de validación española en población oncológica**

Instrumento original	Estudio de validación española	Prueba de referencia	Punto de corte	Sensibilidad	Especificidad	AUC (IC 95 %) <sup>1</sup>	VPP	VPN	CUI+	CUI-	Valor global de la prueba
Hospital Anxiety and Depression Scale (HADS; Zigmond y Snaith, 1983)  Escala Hospitalaria de Ansiedad y Depresión	López-Roig, Terol, Pastor, Neipp, Massutí, RodríguezMarín, Leyda, Martín-Aragón y Sitges, 2000	Para la subescala de ansiedad, el Inventario de Ansiedad Estado-Rasgo (subescala de ansiedad-estado) (STAI-S; Spielberg, 1972; ed. Tea, 1982). Punto de corte: 31 para los varones y 36 para las mujeres.  Para la subescala de depresión, la adaptación española de la Escala de depresión de Zung (SDS; Conde, Escribá e Izquierdo, 1970). Punto de corte: 52.	HADS-D ≥ 4  HADS-A ≥ 9	HADS-D = 0,80  HADS-A = 0,81	HADS-D = 0,81  HADS-A = 0,77	-  -	-  -	-  -	-  -	-  -	-  -
	Costa Requena, Pérez Martín, Salamero Baró y Gil Moncayo, 2009	Entrevista clínica estructurada para los trastornos del eje I del DSM-IV (SCID-I; First, Spitzer, Gibbon y Williams, 1999), empleándose de la siguiente manera:  Diagnóstico de trastorno psiquiátrico (para malestar emocional – puntuación total en la HADS).  Diagnóstico de trastorno psiquiátrico donde prevalece sintomatología depresiva (para la HADS-D).  Diagnóstico de trastorno psiquiátrico donde prevalece sintomatología ansiosa (para la HADS-A).	HADS total ≥ 10  HADS – D ≥ 4  HADS – A ≥ 8	HADS total = 0,74  HADS-D = 0,735  HADS-A = 0,714	HADS total = 0,727  HADS-D = 0,615  HADS-A = 0,765	HADS total = 0,81; (0,74 - 0,87)  HADS-D = 0,74; (0,65 - 0,83)  HADS-A = 0,78; (0,66 - 0,90)	HADS total = 0,551  HADS-D = 0,24  HADS-A = 0,231	HADS total = 0,86  HADS-D = 0,931  HADS-A = 0,964	-  -	-  -	-  -
	Martínez, Durá, Andreu, Galdón, Murgui e Ibañez, 2012	Inventario Breve de Síntomas-18 (BSI-18), con tres subescalas (ansiedad, depresión y somatización) y una puntuación total (Índice Global de Severidad, GSI). Se empleó la traducción española publicada por Derogatis (2000), y se consideró únicamente la puntuación total como medida de malestar emocional.  Las puntuaciones brutas del BSI-18 se convierten en puntuaciones T estandarizadas. Las personas con una puntuación T ≥ 63 puntos en la puntuación total o en, al menos, dos de las tres subescalas, se consideran como casos.	HADS total ≥ 16	HADS total = 0,80	HADS total = 0,92	HADS total = 0,95; (0,94 - 0,97)	-	-	-	-	-

<sup>1</sup> AUC ≥ 0,9 muestra una capacidad discriminativa del test excelente; 0,9 > AUC ≥ 0,8 se considera buena; 0,8 > AUC ≥ 0,7 es aceptable; 0,7 > AUC ≥ 0,6 se considera mediocre; 0,6 > AUC > 0,5 sería fallida (126).

**Abreviaturas:** Subescala de ansiedad de la HADS (HADS-A); Subescala de depresión de la HADS (HADS-D); Área bajo la curva (AUC); Intervalo de confianza (IC); Valor predictivo positivo (VPP); Valor predictivo negativo (VPN); Utilidad clínica positiva (CUI+); Utilidad clínica negativa (CUI-).

**Fuente:** Elaboración propia.

**Tabla 18. Precisión diagnóstica del DT, según los estudios de validación española en población oncológica**

Instrumento original	Estudio de validación española	Prueba de referencia	Punto de corte	Sensibilidad	Especificidad	AUC (IC 95 %) <sup>1</sup>	VPP	VPN	CUI+	CUI-	Valor global de la prueba	
Distress Thermometer (DT; Roth, Kornblith, Batel-Copel, Peabody, Scher y Holland, 1998)	Martínez, Galdón, Andreu e Ibáñez, 2013	BSI-18, con tres subescalas (ansiedad, depresión y somatización) y una puntuación total (GSI) en malestar emocional.  Usando los datos estandarizados por sexo propuestos por Derogatis (2000), las puntuaciones se transforman en puntuaciones T para identificar los casos de malestar (T ≥ 63 en el GSI o, al menos, en dos subescalas).	≥ 5	0,90	0,64	0,82 (0,79 - 0,85)	0,35	0,97	0,32	0,62	68 %	
	Priede, Ruiz, Hoyuela, Herrán y GonzálezBlanch, 2014	HADS. Utilizan la versión española validada por Herrero, Blanch, Perí, De Pablo, Pintor y Bulbena (2003). Consta de 14 ítems, divididos en dos subescalas de depresión y ansiedad. Se empleó la puntuación 15 como punto de corte para discriminar los casos.	≥ 4	0,89	0,56	0,79 (0,71 - 0,86)	0,43	0,93	-	-	-	
		Combinación de la clasificación realizada por la HADS junto con una entrevista diagnóstica semiestructurada basada en los criterios diagnósticos DSM-IV, considerando como casos solamente a aquellos pacientes detectados de modo convergente por ambas pruebas.	≥ 6	0,82	0,76	0,86 (0,79 - 0,93)	0,40	0,95	-	-	-	
	Termómetro de distrés	Martínez, Andreu, Galdón e Ibáñez, 2015	BSI-18, con tres subescalas (ansiedad, depresión y somatización) y una puntuación total (GSI) en malestar emocional.  Usando los datos estandarizados por sexo propuestos por Derogatis (2000), las puntuaciones se transforman en puntuaciones T para identificar los casos de malestar (T ≥ 63 en el GSI o, al menos, en dos subescalas).	DT ≥ 5	DT = 0,93	DT = 0,67	DT = 0,84 (0,80 - 0,87)	DT = 0,25	DT = 0,99	DT = 0,23	DT = 0,66	-
				DIT ≥ 9	DIT = 0,83	DIT = 0,71	DIT = 0,83 (0,79 - 0,87)	DIT = 0,26	DIT = 0,97	DIT = 0,22	DIT = 0,69	-
DT ≥ 5 & IT ≥ 5				DT & IT = 0,74	DT & IT = 0,79	DT = 0,84 (0,80 - 0,87); IT = 0,77 (0,72 - 0,81)	DT & IT = 0,30	DT & IT = 0,96	DT & IT = 0,22	DT & IT = 0,76	-	
Andreu, Galdón, Martínez y Martínez, 2015	BSI-18, con tres subescalas (ansiedad, depresión y somatización) y una puntuación total (GSI) en malestar emocional.  Se consideran casos con malestar clínicamente significativo aquellos con una puntuación T ≥ 63 en el GSI o, al menos, en dos de las subescalas.	DT ≥ 5	DT = 0,94	DT = 0,48	DT = 0,82 (0,73 - 0,91)	DT = 0,42	DT = 0,93	DT = 0,39	DT = 0,45	-		
		DIT ≥ 9	DIT = 0,82	DIT = 0,62	DIT = 0,77 (0,64 - 0,88)	DIT = 0,35	DIT = 0,90	DIT = 0,29	DIT = 0,56	-		
		DT ≥ 5 & IT ≥ 5	DT & IT = 0,53	DT & IT = 0,91	DT = 0,82 (0,73 - 0,91); IT = 0,67 (0,53 - 0,81)	DT & IT = 0,60	DT & IT = 0,88	DT & IT = 0,32	DT & IT = 0,80	-		
Andreu, Martínez y Galdón, 2019	BSI-18, con tres subescalas (ansiedad, depresión y somatización) y una puntuación total (GSI) en malestar emocional.  Usando los datos estandarizados por sexo propuestos por Derogatis (2001), las puntuaciones se transforman en puntuaciones T para identificar los casos de malestar (T ≥ 63 en el GSI o, al menos, en dos subescalas).	≥ 5	0,89	0,59	0,81 (0,74 - 0,87)	0,29	0,97	0,26	0,57	-		

<sup>1</sup> AUC ≥ 0,9 muestra una capacidad discriminativa del test excelente; 0,9 > AUC ≥ 0,8 se considera buena; 0,8 > AUC ≥ 0,7 es aceptable; 0,7 > AUC ≥ 0,6 se considera mediocre; 0,6 > AUC > 0,5 sería fallida (126).

**Abreviaturas:** Termómetro de distrés (DT); Termómetro de impacto (IT); Puntuación combinada de DT e IT (DIT); Inventario Breve de Síntomas-18 (BSI-18); Índice Global de Severidad (GSI); Escala Hospitalaria de Ansiedad y Depresión (HADS); Manual diagnóstico y estadístico de los trastornos mentales, 4ª ed. (DSM-IV); Área bajo la curva (AUC); Intervalo de confianza (IC); Valor predictivo positivo (VPP); Valor predictivo negativo (VPN); Utilidad clínica positiva (CUI+); Utilidad clínica negativa (CUI-).

**Fuente:** Elaboración propia.

En la validación del Cuestionario de Detección de Malestar Emocional (102), se emplearon tres pruebas de referencia y, tras el análisis del área bajo la curva ROC, se decidió determinar el punto de corte  $\geq 9$ .

**Tabla 19. Precisión diagnóstica del DME, según el estudio de validación española en población oncológica**

Instrumento original	Estudio de validación española	Prueba de referencia	Punto de corte	Sensibilidad	Especificidad	AUC (IC 95 %) <sup>1</sup>	VPP	VPN	CUI+	CUI-	Valor global de la prueba
Cuestionario de Detección de Malestar Emocional (DME; Maté, Mateo, Bayés, Bernaus, Casas, González-Barboteo, Limonero, López-Postigo, Sirgo y Viel, 2009)	Limonero, Mateo, Maté-Méndez, González-Barboteo, Bayés, Bernaus, Casas, López, Sirgo y Viel, 2012	Entrevista clínica semiestructurada realizada de forma independiente por un psicólogo o una psicóloga con experiencia en cuidados paliativos. A través de ella, se evalúa globalmente el malestar emocional del paciente con una escala tipo Likert de 5 puntos, desde 1 (Ningún malestar) a 5 (Malestar emocional máximo). El punto de corte establecido fue mayor o igual a 3.	$\geq 9$	0,788 (para punto de corte 9,25)	0,757 (para punto de corte 9,25)	0,81 (0,72 - 0,9)  (para punto de corte 9,25)	0,80	0,782	-	-	-
		Termómetro de distrés (DT). Escala visual numérica de un solo ítem que puntúa de 0 a 10 y mide la valoración global de malestar emocional que experimenta la o el paciente. Punto de corte mayor o igual a 4. No se especifica qué versión española emplea.		0,75 (para punto de corte 9,75)	0,788 (para punto de corte 9,75)	0,85 (0,77 - 0,94)  (para punto de corte 9,75)	0,808	0,792	-	-	-
		Escala Hospitalaria de Ansiedad y Depresión (HADS). Consta de 14 ítems, divididos en dos subescalas de depresión y ansiedad. La suma de ambas subescalas proporciona una medida de malestar global. Se emplea la versión española que usan Tejero, Guimerá, Farré y Peri (2012). Seleccionan el punto de corte $\geq 16$ , siguiendo a Martínez, Porta, Espinosa y Gómez, 2006.		0,909 (para punto de corte 9,25)	0,727 (para punto de corte 9,25)	0,92 (0,84 - 0,99)  (para punto de corte 9,25)	1	0,687	-	-	-

<sup>1</sup> AUC  $\geq 0,9$  muestra una capacidad discriminativa del test excelente;  $0,9 > \text{AUC} \geq 0,8$  se considera buena;  $0,8 > \text{AUC} \geq 0,7$  es aceptable;  $0,7 > \text{AUC} \geq 0,6$  se considera mediocre;  $0,6 > \text{AUC} > 0,5$  sería fallida (126).

**Abreviaturas:** Área bajo la curva (AUC); Intervalo de confianza (IC); Valor predictivo positivo (VPP); Valor predictivo negativo (VPN); Utilidad clínica positiva (CUI+); Utilidad clínica negativa (CUI-).

**Fuente:** Elaboración propia.

Al respecto de la validación del Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de Vida (103), se empleó una combinación de cuatro pruebas de referencia. El análisis del área bajo la curva ROC informa de una buena capacidad discriminativa de la herramienta, y se selecciona el punto de corte  $\geq 28$ .

**Tabla 20. Precisión diagnóstica del ENP-E, según el estudio de validación española en población oncológica**

Instrumento original	Estudio de validación española	Prueba de referencia	Punto de corte	Sensibilidad	Especificidad	AUC (IC 95 %) <sup>1</sup>	VPP	VPN	CUI+	CUI-	Valor global de la prueba
Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida (ENP-E)	Mateo-Ortega, Limonero, Maté-Méndez, Beas, GonzálezBarboteo, Barbero, Ela y GómezBatiste, 2019	La medida criterio es una combinación de: la Escala Hospitalaria de Ansiedad y Depresión (HADS > 17), junto con la puntuación del ítem de calidad de vida del EORTC QLQ-C15-PAL ( $\leq 3$ ), la puntuación del DT (> 5) y el grado de sufrimiento (percepción a largo plazo). Utilizan la versión española de Tejero <i>et al.</i> (112) de la HADS. De las otras escalas no especifican qué versión española emplean. Esta combinación fue construida para indicar la necesidad de tratamiento psicosocial especializado. No queda claro cómo puntúa el grado de sufrimiento y a partir de qué puntuación se considera positivo.	$\geq 28$ (sección A)  Además, algunos ítems (1, 2, 4, 10 y 12) se consideran preguntas clave, de manera que puntuaciones $\geq 4$ en alguno de estos ítems se valoran como posibles signos de alerta de necesidades psicosociales y espirituales.	0,746	0,744	0,82 (0,767 - 0,897)	-	-	-	-	-

<sup>1</sup> AUC  $\geq 0,9$  muestra una capacidad discriminativa del test excelente;  $0,9 > \text{AUC} \geq 0,8$  se considera buena;  $0,8 > \text{AUC} \geq 0,7$  es aceptable;  $0,7 > \text{AUC} \geq 0,6$  se considera mediocre;  $0,6 > \text{AUC} > 0,5$  sería fallida (126).

**Abreviaturas:** European Organization for Research and Treatment of Cancer (EORTC); Quality of Life Questionnaire Core 15 Palliative (QLQ-C15-PAL); Área bajo la curva (AUC); Intervalo de confianza (IC); Valor predictivo positivo (VPP); Valor predictivo negativo (VPN); Utilidad clínica positiva (CUI+); Utilidad clínica negativa (CUI-).

**Fuente:** Elaboración propia.

En síntesis, la capacidad discriminativa de cualquiera de las herramientas de detección precoz analizadas es buena, tomando el AUC valores superiores a 0,80 para las EVA de ansiedad y depresión, para el DT, el DME, la HADS y el ENP-E. Para el BSI-18 y el ESAS-r, este índice no se calculó por falta de prueba de referencia.

#### 4.2.3. Valoración global de las escalas

Complementariamente, se llevó a cabo una valoración global para cada una de las escalas siguiendo el procedimiento propuesto por Vodermaier *et al.* (48). En la siguiente tabla, se muestran los resultados de esta valoración.

**Tabla 21. Valoración global de las herramientas, según el procedimiento de Vodermaier *et al.* (48)**

Longitud de la escala	Medida	N.º de ítems	N.º de estudios	N.º de participantes	Fiabilidad <sup>1</sup>	Medida criterio <sup>2</sup>	Validez <sup>3</sup>	Valoración
Medidas ultracortas	DME	2	1	105	Moderada	Alta	Moderada	Moderada
	DT	1	5	1799	-	Moderada	Moderada	Moderada
	EVA de an y de	2	1	53	-	Moderada	Alta	Pobre
Medidas cortas	BSI-18	18	3	2135	Alta	-	-	Pobre
	ENP-E	13	1	129	Moderada	Moderada	Moderada	Aceptable
	ESAS-r	10	1	66	Alta	-	-	Pobre
	HADS	14	3	1610	Alta	Moderada	Moderada	Moderada

<sup>1</sup> Baja = Alfa de Cronbach < 0,60; Moderada = Alfa de Cronbach ≥ 0,60 y < 0,80; Alta = Alfa de Cronbach ≥ 0,80. Cuando la información sobre la fiabilidad era baja o ausente, la escala se valoraba como pobre, a menos que la fiabilidad no fuese aplicable (en caso de medidas de uno o dos ítems).

<sup>2</sup> Baja = diagnóstico clínico; Moderada = cuestionario validado; Alta = entrevista clínica estructurada (criterio estándar).

<sup>3</sup> Baja = sensibilidad y especificidad promedio < 0,60; Moderada = sensibilidad y especificidad promedio ≤ 0,60 y < 0,80; Alta = sensibilidad y especificidad promedio ≥ 0,80.

**Abreviaturas:** DME = Cuestionario de Detección de Malestar Emocional; DT = Termómetro de distrés; EVA de an y de = Escalas Visuales Analógicas de ansiedad y depresión; BSI-18 = Inventario Breve de Síntomas18; ENP-E = Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida; ESAS-r = Cuestionario de evaluación de síntomas de Edmonton, versión revisada; HADS = Escala Hospitalaria de Ansiedad y Depresión.

**Fuente:** Elaboración propia.

Las escalas BSI-18 y ESAS-r reciben una valoración global pobre, dado que los estudios de validación (90, 92, 93, 100) no contemplan una prueba de referencia o criterio con que explorar su sensibilidad y especificidad. En el caso de las EVA de ansiedad y depresión, la valoración global también resulta pobre, debido al reducido número de participantes en el estudio (n = 53) (101).

Por su parte, tanto el Termómetro de distrés, como la HADS y el cuestionario DME reciben una valoración global moderada, y el ENP-E, aceptable.

#### 4.2.4. Síntesis global de la evidencia

Para finalizar, con la idea de facilitar una visión holística de la evidencia analizada, se elaboró la tabla 22. Esta tabla resume las características y los resultados más relevantes de cada estudio, detallados en los apartados anteriores, así como la valoración tanto por estudio como por escala.

**Tabla 22. Resumen de las características más relevantes de los estudios de validación y las herramientas de detección precoz analizados**

Longitud de la escala	Herramienta	Estudio	Muestra	Punto de corte	Consistencia interna	Auc (ic 95 %)	Sensibilidad	Especificidad	Valoración estudio: Quadas-2 <sup>1</sup>	Valoración global escala
ULTRACORTA	EVA de an y de (2 ítems)	Martínez, 2006	53 (paliativos)	≥ 10	-	0,870	0,80	0,939	☺	Pobre
	DT (1 ítem)	Martínez, 2013	962	≥ 5	-	0,82 (0,79 - 0,85)	0,90	0,64	☺	Moderada
		Priede, 2014	135	≥ 6	-	0,86 (0,79 - 0,93)	0,82	0,76	?☺☺	
		Martínez, 2015	385	≥ 5	-	0,84 (0,80 - 0,87)	0,93	0,67	☺	
		Andreu, 2015	81	≥ 5	-	0,82 (0,73 - 0,91)	0,94	0,48	☺	
		Andreu, 2019	236	≥ 5	-	0,81 (0,74 - 0,87)	0,89	0,59	☺	
DME (2 ítems)	Limonero, 2012	105 (paliativos)	≥ 9	0,685	0,81 (0,72 - 0,9)	0,788	0,757	☺☺	Moderada	
CORTA	BSI-18 (18 ítems)	Galdón, 2008	175	T ≥ 63	0,82	-	-	-	?☺	Pobre
		Martínez, 2019	1183		0,91	-	-	-	☺	
		Calderon, 2020	877		0,91	-	-	-	☺	
	ESAS-r (10 ítems)	Carvajal, 2013	66 (avanzado)	¿?	0,86	-	-	-	?☺☺	Pobre
	HADS (14 ítems)	López-Roig, 2000	484	HADS-D ≥ 4 HADS-A ≥ 9	α HADS-A = 0,85 α HADS-D = 0,87	-	HADS-D = 0,80 HADS-A = 0,81	HADS-D = 0,81 HADS-A = 0,77	?☺☺	Moderada
		Costa, 2009	234	HADS t ≥ 10 HADS-D ≥ 4 HADS-A ≥ 8	α HADS-A = 0,82 α HADS-D = 0,84	HADS t = 0,81; (0,74 - 0,87) HADS-D = 0,74; (0,65 - 0,83) HADS-A = 0,78; (0,66 - 0,90)	HADS total = 0,74 HADS-D = 0,735 HADS-A = 0,714	HADS total = 0,727 HADS-D = 0,615 HADS-A = 0,765	?☺	
		Martínez, 2012	892	HADS t ≥ 16	α HADS t = 0,89 α HADS-A = 0,83 α HADS-D = 0,83	0,95; (0,94 - 0,97)	HADS total = 0,80	HADS total = 0,92	☺	
ENP-E (13 ítems)	Mateo-Ortega, 2019	129 (paliativos)	≥ 28	0,75	0,82 (0,767 - 0,897)	0,746	0,744	?☺	Aceptable	

<sup>1</sup> ☺ = No genera preocupaciones respecto al riesgo de sesgo ni a la aplicabilidad; ☺☺ = Algunos dominios generan preocupaciones respecto al riesgo de sesgo o a la aplicabilidad; ?☺ Algunos dominios generan incertidumbre respecto al riesgo de sesgo o a la aplicabilidad; ?☺☺ = Algunos dominios generan preocupaciones e incertidumbre respecto al riesgo de sesgo o a la aplicabilidad de los resultados.

**Abreviaturas:** DME = Cuestionario de Detección de Malestar Emocional; DT = Termómetro de distrés; EVA de an y de = Escala Visual Analógica de Ansiedad y Depresión; BSI-18 = Inventario Breve de Síntomas 18; ENP-E = Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida; ESAS-r = Cuestionario de evaluación de síntomas de Edmonton, versión revisada; HADS = Escala Hospitalaria de Ansiedad y Depresión.

**Fuente:** Elaboración propia.

## 5. Consideraciones de implementación

### 5.1. Aspectos organizativos

**¿La introducción y el uso de una herramienta de detección precoz del malestar emocional en población oncológica adulta requiere cambios relevantes en la organización y la prestación de servicios?**

La introducción de herramientas de detección precoz del malestar emocional podría conllevar cambios organizativos relevantes, sobre todo si se incorpora de forma sistemática o como parte de un programa de cribado en pacientes con cáncer.

Se ha descrito que el hecho de realizar un cribado de malestar emocional puede requerir ciertos cambios en el comportamiento del personal clínico y en la cultura organizacional (49). En cuanto al personal implicado, podría ser necesario un coordinador o una coordinadora, además del tiempo del personal para participar en el desarrollo del programa y el tiempo para formarse (62). A este respecto, la GPC canadiense recomienda crear un grupo coordinador que lidere la implementación del cribado de malestar, y que incorpore a los diferentes agentes de interés (equipo multidisciplinar, pacientes y personal gestor) (14).

Por su parte, el panel de la NCCN® ha identificado algunos principios de implementación para guiar a las instituciones en el desarrollo de un protocolo de cribado de malestar emocional y de derivación apropiada y seguimiento (4). Estos principios comprenden la creación de un grupo de trabajo interdisciplinar, que incluya personal de la medicina, la enfermería, la psicología, las tecnologías de la información, el trabajo social, capellanes y personal administrativo; el apoyo explícito y el respaldo de los y las líderes institucionales; el desarrollo y la ejecución de un programa piloto previo a la implementación a gran escala; la consideración de los recursos existentes y el flujo de trabajo y proceso actuales (4). Además, ha de determinarse la frecuencia de cribado, el algoritmo de respuesta, la estrategia de comunicación de los resultados al equipo y de incorporación de los resultados a la historia clínica del paciente, y los puntos de corte establecidos para las respectivas derivaciones (4). En este sentido, la NCCN® subraya la necesidad de contar con profesionales de la salud mental a quienes se pueda acudir con facilidad para ofrecer la atención psicosocial que sea necesaria en cada caso, bien como miembros del personal del propio equipo de oncología o bien por derivación a un equipo de salud mental (4).

En el caso del contexto español, no se han identificado GPC aquí desarrolladas que aborden la cuestión organizativa, pero cabría resaltar que la Comisión de Test del Consejo General de la Psicología de España (COP) expone que el uso de los test se ha de restringir únicamente a los y las profesionales que cuenten con la cualificación adecuada para ello (127). Esto significa que la implementación del uso de un test psicológico en un determinado programa de cribado debería estar supervisada y validada por un psicólogo o una psicóloga, independientemente de qué profesional en concreto lo administre, pues

sí que pueden delegarse funciones en otros miembros del equipo multidisciplinar del que forme parte. Es decir, sería necesario que, entre los miembros del equipo encargado de llevar a cabo el programa de cribado, se cuente con la figura de un o una profesional de la psicología. Esto sería aplicable tanto en unidades oncológicas de atención especializada como en Unidades de Consejo Genético en Cáncer Hereditario.

Asimismo, como se ha subrayado en las GPC consultadas (4, 14, 31, 51), resulta de especial relevancia no solo que exista un programa de cribado en el que se establezca el uso sistemático de una herramienta de detección precoz del malestar emocional, sino que también garantice un circuito de derivación que pueda dar respuesta a tiempo a las necesidades psicosociales de los y las pacientes.

La literatura científica ha puesto de relieve que todos estos aspectos impactan en las probabilidades de que la implementación del cribado de distrés emocional en población oncológica resulte exitosa o no (128).

## 5.2. Aspectos éticos

### **¿La introducción de una herramienta de detección precoz del malestar emocional en población oncológica adulta supone algún conflicto ético relevante?**

No se han descrito conflictos éticos relevantes con respecto al empleo de herramientas de detección precoz del malestar emocional en pacientes con cáncer ni en las guías de práctica clínica analizadas ni en los estudios primarios.

En el caso de que la herramienta se emplee como parte de un programa de cribado, el programa debe cumplir el requerimiento ético de que el beneficio supere los posibles riesgos (40).

Asimismo, han de tenerse en cuenta otras consideraciones éticas como la equidad en el acceso, el respeto a la autonomía y el derecho a la intimidad y la confidencialidad a la hora de implementar dicho programa (40). De manera general, la justificación ética de los programas de cribado se ha de basar en los siguientes aspectos: la prueba de cribado es adecuada, precisa y fiable; el diagnóstico precoz produce un beneficio para el o la paciente, la familia y la sociedad; el beneficio del cribado está razonablemente equilibrado frente a los costes financieros y de otro tipo; hay un proceso organizado de diagnóstico, consejo, tratamiento y seguimiento de las y los pacientes identificados mediante la prueba de cribado; el funcionamiento del programa se monitoriza a lo largo del tiempo para garantizar que se cumplen los estándares de calidad y se alcanzan los objetivos propuestos (129).

Estos aspectos éticos han de ser extensibles al uso no sistemático o puntual de las herramientas —fuera de un programa estructurado, realizado por iniciativa de un o una profesional o a petición de los y las pacientes—, que también debería guiarse por ese mismo balance riesgo-beneficio, garantizando la equidad, la autonomía y el derecho a la intimidad y la confidencialidad.

Además, el empleo de estas herramientas y las intervenciones derivadas de la detección del malestar emocional —cita de seguimiento, abordaje en la propia consulta, derivación a otros dispositivos especializados, etc.— deberán estar guiados por un proceso de toma de decisiones compartidas.

### 5.3. Aspectos sociales y relacionados con la visión de los y las pacientes

#### **¿La introducción de una herramienta de detección precoz del malestar emocional en población oncológica adulta supone algún conflicto relacionado con la visión de los y las pacientes relevante?**

La guía NCCN® enfatiza que una de las barreras para recibir tratamiento psicosocial es la existencia de estigma a la hora de comunicar problemas psicológicos, un aspecto que influye de manera bidireccional: haciendo que los y las pacientes con cáncer a menudo no informen sobre su malestar al equipo médico, y que las y los profesionales no pregunten sobre estas cuestiones a sus pacientes (4). Por ello, el empleo de escalas estandarizadas podría ayudar a vencer esta barrera.

Con respecto al propio constructo a medir, en la atención oncológica se considera el malestar como la medida emocional clave informada por el o la paciente (128). El concepto de malestar emocional tiene la ventaja de ser menos estigmatizador que otros conceptos —como la depresión—, cuenta con una mayor aceptabilidad por parte de los y las pacientes (128), produce menos vergüenza y puede medirse por autoinforme (4).

De igual manera, se deben considerar y prever los aspectos que puedan influir en la aceptabilidad de la prueba por parte de la población cribada (40), pues cualquier beneficio potencial de un método de cribado estará limitado por el nivel de aceptación del procedimiento entre las y los pacientes (130). De nada serviría desarrollar un instrumento excelente para la detección precoz del distrés en cuanto a bondad psicométrica si los y las pacientes no lo aprueban (130). En esta línea, los estudios primarios incluidos en este informe apenas han incorporado información sobre la satisfacción o la aceptabilidad de las herramientas de detección del malestar emocional evaluadas. Uno de los estudios de validación de la HADS (96) alude, en la introducción, a que una de sus características es su aceptabilidad por parte de los y las pacientes —pero esta afirmación no se apoya en ninguna cita bibliográfica—. La aceptabilidad de la herramienta también es una característica que se asume en el apartado introductorio de dos de los estudios de validación del DT (98, 99) —aunque tampoco aportan ninguna cita bibliográfica que avale este hecho—. En otros dos estudios de validación del Termómetro de distrés se subraya, en el apartado de discusión, la importancia de que futuras investigaciones sobre el DT incluyan el estudio de la aceptabilidad de su uso tanto por parte de profesionales de la salud como de pacientes, dado que la factibilidad de un cribado rutinario de distrés precisa de que las personas implicadas consideren que las herramientas de cribado son útiles y satisfactorias en la práctica clínica habitual (35, 97). No se ha encontrado información al respecto en las GPC analizadas.

Por otra parte, en el marco general de construcción y validación de test y escalas, existen unas primeras fases en las que se atiende a su precisión diagnóstica y se estudian, principalmente, las propiedades psicométricas (130), y unas fases posteriores en las que adquieren mayor relevancia las cuestiones relacionadas con la implementación (130). Los estudios de validación recogidos en este informe se enmarcan en las primeras. En estos estadios, la visión de las y los pacientes suele incluirse en la fase de construcción y pilotaje (131). La finalidad de cualquier estudio piloto es evaluar el funcionamiento general del instrumento de medida en una muestra de participantes con características similares a la población objeto de interés (131). Esta fase permite detectar, evitar y corregir posibles errores, conocer las opiniones de los y las pacientes acerca de cuestiones técnicas como la legibilidad del test o el tamaño de la fuente, y cuestiones más relacionadas con el contenido como el marco temporal, la escala de respuesta o el grado de comprensión de los ítems.

A este respecto, la mayoría de los estudios de validación analizados en este informe no describen la fase de estudio piloto de la herramienta, pues la validación en población oncológica española toma como punto de partida una versión española de la herramienta ya validada previamente (como en el caso del BSI-18, la HADS o el DT). Solo se describe una fase piloto en dos estudios de validación (100, 103). En concreto, en el estudio de validación del ESAS-r (100) se realiza un estudio piloto con 14 pacientes oncológicos, cuya opinión se tuvo en cuenta para llevar a cabo algunos cambios en el formato, de manera que el diseño de la herramienta fuese más claro. Posteriormente, en el propio estudio de validación también se comprobó que los y las pacientes preferían la nueva versión de la herramienta (ESAS-r) frente a la antigua (ESAS) (100). En el estudio de validación del ENP-E (103), se presenta el establecimiento de un plan de incorporación de la visión de los y las pacientes para evaluar la validez aparente del instrumento. Se solicitó a 20 pacientes que calificaran las preguntas que conforman la herramienta, utilizando una escala tipo Likert de 6 puntos (Excelente, bueno, regular, malo, muy malo, innecesario) en cuanto a: (1) claridad de la redacción; (2) adecuación de las opciones de respuesta; (3) idoneidad y pertinencia de las preguntas para otros pacientes en situaciones similares. También podían añadir observaciones en cada pregunta. Este feedback se contempló para mejorar las preguntas del test. Por su parte, no se ha identificado información a este respecto en las GPC analizadas.

Para concluir, cabe considerar que, cuando se use la herramienta en la práctica clínica, se deberá garantizar que la persona que cumplimenta el cuestionario lo entienda y facilitar la información y el apoyo que sean necesarios.

## 6. Discusión

### 6.1. Discusión de la metodología

Para desarrollar este informe se siguió la metodología planteada para la elaboración de informes de evaluación de tecnologías sanitarias en la RedETS (87).

Con el fin de asegurar la pertinencia de los resultados y de evitar pérdidas de información en el proceso de búsqueda, se diseñó una estrategia específica que combinase términos libres y descriptores, agrupando todos los sinónimos relevantes para nuestro objetivo, y se adaptó para interrogar las bases de datos específicas de guías de práctica clínica y generales recomendadas en los manuales principales (132). También se ejecutó una búsqueda específica para recuperar literatura gris, y no se aplicaron límites temporales o por idioma. De este modo, la búsqueda de bibliografía no se considera una limitación. Debido a la exhaustividad en la estrategia de búsqueda planteada, el riesgo de haber perdido estudios relevantes es mínimo.

Las GPC se emplearon para los dominios de descripción del malestar emocional, las herramientas disponibles y para describir las consideraciones de implementación. Para la síntesis de evidencia de los estudios primarios —estudios de validación de herramientas—, se utilizaron las tablas de evidencia resultantes de cubrir las fichas de lectura crítica para pruebas diagnósticas de la plataforma FLC 3.0 (80). No obstante, hubo que adaptar las tablas finales *ad hoc* a esta revisión, pues no se recogían ciertos índices relevantes para analizar las propiedades psicométricas de las herramientas de detección precoz en el ámbito psicológico —por ejemplo, la consistencia interna o los índices de bondad de ajuste del análisis factorial—. Asimismo, la variabilidad en el tipo de análisis y los índices informados por parte de cada uno de los estudios hizo que la comparación entre estudios representase un desafío. Las investigadoras trataron de contrarrestar estas dificultades realizando la tarea de síntesis por pares, y centrando el análisis en los resultados más relevantes. Dada la heterogeneidad de los diseños, las herramientas y los índices evaluados, se ha desestimado efectuar una síntesis cuantitativa.

Con respecto a la valoración de la calidad de la evidencia, se empleó la herramienta QUADAS-2 para todos los estudios de validación. Sin embargo, es relevante destacar que los estudios de validación del BSI-18 y del ESAS-r analizaron ciertas propiedades psicométricas de las herramientas, pero no su precisión diagnóstica frente a una prueba de referencia, por lo que este aspecto no ha podido considerarse en dichos casos. De forma complementaria, se valoró cada una de las herramientas siguiendo el método empleado por Vodermaier *et al.* (48) en su revisión sistemática. Este método aporta una valoración global sobre cada una de las herramientas analizadas, valorando el tamaño de las muestras de validación, la consistencia interna, la sensibilidad y la especificidad de la escala, y el tipo de medida criterio empleada —siendo esta última valorada en función de si consistía en un diagnóstico clínico, un cuestionario validado o una entrevista clínica estructurada—. Con todo, no existe una valoración de la calidad de esa medida criterio en sí, lo que hace que, por

ejemplo, cuando se emplea el BSI-18 como medida criterio para una validación, no exista una penalización en la valoración, pese a no contar con validaciones completas publicadas de dicha herramienta en nuestro idioma. Aunque en el protocolo se había planteado recurrir a otras herramientas, de ser necesario, como el AMSTAR-2 o las directrices para la traducción y adaptación de los test (2ª edición) de la International Commission Test, no se empleó ninguna de ellas, ya que no se localizaron revisiones sistemáticas, ni metanálisis, ni se describe con detalle el proceso de adaptación y traducción de aquellas herramientas originalmente desarrolladas en inglés, pues la validación en población oncológica española toma como punto de partida una versión española de la herramienta ya validada y publicada con anterioridad (como en el caso del BSI-18, la HADS o el DT). En el caso de la validación del ESAS-r (100), el proceso de adaptación y traducción se describe de forma muy breve, lo que tampoco permite dicho análisis.

## 6.2. Discusión de los resultados de efectividad

Los resultados obtenidos de esta revisión responden al objetivo de este informe —proporcionar una evaluación de las características y las propiedades psicométricas de las herramientas para la detección precoz del malestar emocional en población oncológica adulta validadas en el contexto español que puedan utilizarse en el contexto del SNS— y a la pregunta de investigación formulada.

A través de la búsqueda bibliográfica, se localizaron un total de 15 estudios de validación (de 7 herramientas distintas). A lo largo de la sección de resultados se han detallado las características de los estudios realizados y la calidad de dicha evidencia, así como las propiedades psicométricas y la precisión diagnóstica de las escalas y su valoración global, con el objetivo de proporcionar toda la información relevante que se debe tener en cuenta para seleccionar una u otra herramienta.

En primer lugar, cabe destacar que, siguiendo la clasificación propuesta por Vodermaier *et al.* (48), en función de la longitud de las herramientas, los instrumentos analizados en este informe se clasifican como medidas ultracortas (entre 1 y 4 ítems) o cortas (entre 5 y 20 ítems). Observando los valores que componen la tabla resumen (tabla 22), los resultados coinciden con lo planteado por autoras como Martínez (130), que apuntan que las medidas ultracortas suelen poseer una menor precisión diagnóstica, pues, aun teniendo niveles de sensibilidad adecuados, los valores de especificidad suelen ser de bajos a moderados. No obstante, estas medidas representan una ventaja en el ámbito sanitario, ya que son más sencillas de implementar dentro de la atención rutinaria (130). El tiempo necesario para completar la prueba, el personal vinculado a su administración y la sencillez para obtener la puntuación total favorecen su aplicación, incluso en los servicios más concurridos. Una solución que se apunta en este sentido pasa por considerar que, aunque las medidas ultracortas puedan no ser suficientes por sí mismas, pueden considerarse como un primer estadio del propio cribado (14, 130).

El empleo de una u otra herramienta también podrá diferir según el contexto y el tipo de pacientes con quienes vaya a utilizarse. Otro criterio para tener en cuenta a la hora de decantarse por una u otra herramienta puede ser la calidad de la evidencia de los estudios

de validación de cada uno de los instrumentos, el número de estudios de validación que respalda cada herramienta, así como la valoración global de cada escala, fruto de los índices estimados de sus propiedades psicométricas y su precisión diagnóstica.

Teniendo todo ello en consideración, encontramos 3 medidas ultracortas, dos para población paliativa (EVA de ansiedad y depresión, y cuestionario DME) y una para población oncológica general (Termómetro de distrés). En el caso de las EVA (101) (con un punto de corte  $\geq 10$ ), la valoración de la calidad del único estudio de validación no genera preocupaciones, aunque la valoración global de la escala resulta pobre dado su reducido tamaño muestral. En cuanto al cuestionario DME (102) (con un punto de corte  $\geq 9$ ), la valoración global de la escala resulta moderada, y la valoración de la calidad del estudio revela riesgo de sesgo respecto del dominio prueba índice. Con relación al DT (con un punto de corte  $\geq 5$ ), es la herramienta que cuenta con un mayor número de estudios de validación (35, 91, 97-99). Su valoración global es moderada, y en cuatro (35, 97-99) de los cinco estudios no se detecta riesgo de sesgo.

Por otra parte, localizamos 4 medidas cortas, dos para población paliativa (ESAS-r y ENP-E) y dos para población oncológica general (BSI-18 y HADS). La utilización del ESAS-r (100) no estaría respaldada ni por la valoración global de la escala (pobre) ni por la valoración de la calidad de la evidencia, pues su único estudio de validación suscita preocupaciones tanto a nivel de riesgo de sesgo como de aplicabilidad (no se explicita el punto de corte). Con respecto al ENP-E (103), con un punto de corte  $\geq 28$ , la valoración global resulta aceptable, mientras que la valoración de la calidad de la evidencia sugiere riesgo de sesgo incierto en la selección de pacientes y la prueba de referencia. En relación con el BSI-18, la calidad de los tres estudios no plantea apenas preocupaciones sobre riesgo de sesgo y aplicabilidad, pero su valoración global resulta pobre, ya que los estudios no estiman su precisión diagnóstica frente a una prueba de referencia (90, 92, 93). Con referencia a la HADS, la valoración global se considera moderada, mientras que la calidad de la evidencia de los estudios presenta riesgos de sesgo en dos (94, 95) de los tres estudios de validación. En cuanto a la puntuación total de corte recomendada en el caso de la HADS, no existe consenso, dado que uno de los estudios solo aporta un punto de corte por cada subescala pero no de manera global (94), otro estudio propone el punto de corte 10 (95) y el otro propone como punto óptimo  $\geq 16$  (96), aun aceptando cualquier punto de corte entre 13 y 16 por sus adecuados niveles de sensibilidad y especificidad asociados. La heterogeneidad en el punto de corte que señala cada estudio de validación de la HADS no es un dilema nuevo, tal y como subraya el equipo autor de una revisión bibliográfica de estudios de la Escala Hospitalaria de Ansiedad y Depresión en muestras españolas (133) publicada en 2015. El equipo autor de la revisión resuelve la falta de consenso sugiriendo un punto de corte  $\geq 13$ , en concordancia con lo que apuntan estudios anteriores (134, 135). A la luz de los resultados arrojados por los tres estudios analizados en este informe, también parece razonable aceptar ese punto de corte.

A modo de resumen, podría decirse que destaca el Termómetro de distrés (con un punto de corte  $\geq 5$ ) como herramienta de detección precoz del malestar emocional, por su brevedad (1 ítem), por haber sido validado en cinco estudios —de los cuales cuatro cuentan con un riesgo de sesgo bajo— cuyas muestras provienen de población oncológica

en diferentes estadios de la enfermedad, así como por contar con una valoración global de la escala moderada. Asociada a este test, se encuentra la Lista de Problemas, que puede utilizarse de forma complementaria para identificar el tipo de problemas que le preocupan al paciente y, así, orientar la posible derivación a otros y otras profesionales. No obstante, ha de tenerse en cuenta que esta lista no proporciona mejoras en la sensibilidad o la especificidad del cribado, sino únicamente información cualitativa.

Para refinar el proceso de detección precoz del malestar emocional en pacientes con cáncer y superar los inconvenientes derivados de emplear una medida de un solo ítem, el uso del Termómetro de distrés podría emplearse en un primer estadio del cribado, y combinarse posteriormente con una escala de mayor longitud, lo que contribuiría a que la especificidad del cribado en su conjunto mejorase. Teniendo en cuenta la valoración de las herramientas con respecto al baremo de Vodermaier (48), así como el número de estudios de validación y la valoración de la calidad de la evidencia de estos, destacaría la escala HADS para este fin. La utilización de una prueba más larga en el segundo estadio del cribado comporta una doble ventaja: por una parte, supone una mejora con respecto a la precisión diagnóstica del cribado en su conjunto y, por otra, permite obtener información cualitativa más detallada sobre las fuentes de malestar y, por tanto, orientar la derivación y la intervención de una manera más sencilla, a sabiendas de que existen diferentes subgrupos de pacientes con cáncer: aquellos que no reportan malestar y no necesitan apoyo psicológico; aquellos con altos niveles de malestar clínicamente significativo, que necesitan apoyo psicológico con urgencia, y aquellos con niveles moderados de malestar que necesitan apoyo bien psicológico o bien social o físico (136).

En cualquier caso, la variabilidad en el número y las características de las herramientas no tiene por qué ser vista como un inconveniente, sino que puede entenderse como un abanico de opciones para que seleccione el personal facultativo o la propia administración sanitaria, en función de su objetivo —prefiriendo escalas con un determinado balance de sensibilidad/especificidad u otro según vayan a utilizarse para realizar un cribado rutinario o una identificación de casos clínicos—, el estadio de la enfermedad del paciente —seleccionando una escala validada con población oncológica general o con población paliativa—, o los aspectos organizativos —optando por escalas ultracortas o cortas, según la disponibilidad de tiempo para la administración, la puntuación y la interpretación que tenga el equipo sanitario que la aplique—.

De cara a obtener más y mejores resultados que reafirmen la utilidad de las pruebas para el propósito que persiguen, sería relevante que se amplíe el número de estudios que evalúen de manera completa la calidad psicométrica y la precisión diagnóstica de las herramientas de detección precoz del malestar emocional, ya que el número total de estudios de validación para cada prueba es reducido —incluso únicamente uno para cuatro de las escalas analizadas— y, en algunos casos, resultan incompletos al no incluir prueba de referencia. Sería particularmente interesante que se llevasen a cabo este tipo de estudios para la herramienta BSI-18, pues, a pesar de que otros estudios la emplean como prueba de referencia, no se cuenta con estudios con muestra española oncológica que evalúen su precisión diagnóstica.

### 6.3. Discusión de los aspectos organizativos, éticos, sociales y relacionados con la visión de los y las pacientes

La introducción de herramientas de detección precoz del malestar emocional puede tener un impacto organizativo sustancial, sobre todo si se introduce como parte de un programa de cribado o evaluación sistemática. Por ello, varias guías recomiendan la creación de subgrupos de trabajo que coordinen y lideren su implementación (4, 14). Además de efectuarse el cribado en las unidades oncológicas de atención especializada, cabría pensar que también sería beneficioso implementarlo en las Unidades de Consejo Genético en Cáncer Hereditario, ya que entre sus funciones se encuentra evaluar el estado emocional de la persona usuaria y proporcionar apoyo psicológico en los diferentes momentos del proceso (137, 138).

Una revisión sobre los beneficios del cribado de distrés emocional en población oncológica desarrollada por Mitchel (128) mostró que, en los 24 estudios analizados, las barreras clínicas y organizativas tenían un efecto significativo en el éxito de la implementación del cribado. En el ámbito clínico, los principales obstáculos detectados eran la falta de tiempo, la falta de formación y la falta de habilidades personales o de confianza para llevarlo a cabo (128). A nivel organizativo, las barreras incluían la falta de recursos y la ausencia de una estrategia de cribado (128). De hecho, la barrera principal radicaba en no recibir una atención adecuada tras el cribado, lo que demuestra que los cuidados posteriores son probablemente el paso clave del proceso, y que un resultado clínicamente significativo debería ir seguido de una evaluación clínica y una gestión competente (128). De los 24 estudios analizados, 17 reportaron beneficios significativos del cribado en resultados primarios (bienestar del paciente) o secundarios (número de diagnósticos precisos registrados por el personal facultativo, comunicación personal médicopaciente, derivaciones a servicios especializados o ayuda psicosocial prestada), 6 informaron de ausencia de efecto y 1 encontró un efecto perjudicial no significativo (128). Asimismo, se halló que el cribado era más efectivo si iba vinculado a una intervención o derivación preceptiva (128). Como conclusiones, Mitchell señala que el cribado de distrés emocional en población oncológica tiene el potencial para influenciar el bienestar de los y las pacientes, pero solo si se abordan las barreras mencionadas (128). Complementariamente, otros autores y autoras han indicado que sería recomendable que el cribado se realice en la fase de tratamiento activo —y que no se espere a que finalice—, momento en el que se maximizaría la detección de pacientes con malestar emocional que todavía no han recibido ayuda al respecto (139).

En esta línea, el equipo autor de una revisión sistemática de ensayos clínicos aleatorizados y no aleatorizados que comparaban el efecto de las intervenciones de cribado de bienestar psicosocial y las necesidades de atención de las personas con cáncer en comparación con la atención habitual (140) expone que ningún estudio reportaba eventos adversos, pero que solo algunos estudios mostraban efectos beneficiosos del cribado en el bienestar de los y las pacientes, mientras que otros no encontraban ningún efecto, o incluso algunos hallaban efectos negativos —como mayores niveles de necesidades psicológicas no cubiertas o una satisfacción menor—. Sin embargo, en esta revisión se excluyeron los

estudios que evaluaban el cribado integrado como un elemento en intervenciones más complejas —por ejemplo, la terapia, el coaching, los programas de atención...— (140). Estos resultados refuerzan la necesidad de emplear las herramientas de detección precoz del malestar emocional dentro de un programa estructurado que asegure la atención posterior a las y los pacientes que la necesiten, pues la evidencia analizada no permite asegurar el efecto beneficioso del cribado en sí mismo (140). A su vez, ha de tenerse en cuenta que las barreras que obstaculizan el éxito de la implementación de un programa de cribado también están relacionadas con la visión de los y las pacientes. En este sentido, una revisión sistemática, que dio cuenta de las barreras experimentadas y percibidas por los y las pacientes en la prestación de atención psicosocial a personas adultas con cáncer (141), advirtió que la principal barrera que señalan las y los pacientes es la ausencia de necesidad percibida de este tipo de atención (141). En concreto, algunos estudios apuntan que casi la mitad de las personas que declaraban malestar significativo no había recurrido a los servicios psicosociales ni tenía intención de hacerlo, mientras que otros estudios indicaban que las y los pacientes con niveles de malestar más elevados eran significativamente más propensos a manifestar miedo al impacto negativo de una intervención psicosocial, mientras que aquellos y aquellas con niveles de malestar bajos eran significativamente más propensos a afirmar que no necesitaban apoyo (141).

En cuanto a los aspectos éticos, cualquier evaluación de este tipo deberá garantizar la equidad, la autonomía y la confidencialidad de la información. Asimismo, cabría pensar que uno de los conflictos éticos del uso de herramientas de detección precoz del malestar emocional sería el impacto de su imprecisión diagnóstica en las y los pacientes oncológicos. Tal y como reflexiona Martínez (130), la implementación de este tipo de cribado pudiera dar lugar a la derivación de pacientes que en realidad no presentan niveles clínicamente significativos de distrés (falsos positivos) a servicios especializados de asistencia (130), lo que podría acarrear una utilización innecesaria de recursos de atención psicosocial. Además, los falsos positivos podrían tener la consecuencia no deseada de generar distrés donde antes no lo había (130). No obstante, la precisión generalmente buena de las herramientas de detección precoz para cribar niveles elevados de distrés puede considerarse como un importante contrapunto que supera el inconveniente descrito (130). Es más, dados los dos posibles tipos de errores (falsos positivos y falsos negativos), la no detección de una persona paciente con niveles significativos de distrés tendría consecuencias más graves para su calidad de vida y para la calidad de la atención que lo contrario (130). Las autoras y los autores de los estudios de validación analizados han contemplado este aspecto, puesto que, a la hora de establecer el punto de corte de las herramientas de detección precoz, han primado la sensibilidad frente a la especificidad de los instrumentos.

En lo que concierne a los posibles conflictos sociales y la visión de los y las pacientes, apenas se ha identificado información sobre la aceptabilidad, la satisfacción o las perspectivas de las y los pacientes ni en las GPC recuperadas ni en los estudios analizados. Únicamente, el estudio de validación del ESAS-r (100) informa sobre la facilidad de uso de la herramienta según los y las pacientes; y en el estudio de validación del ENP-E (103), las y los pacientes califican las preguntas que conforman la herramienta en base a la claridad

de la redacción, la adecuación de las opciones de respuesta, la idoneidad y la pertinencia de las preguntas.

Sería esperable que futuras investigaciones, enmarcadas en las fases posteriores del proceso de validación de herramientas, prestaran mayor atención a cuestiones relacionadas con su implementación (130). Por ejemplo, resultaría relevante que la prueba sea aceptable para la población diana, teniendo en cuenta la diversidad social y cultural, y las peculiaridades de grupos desfavorecidos o de población con discapacidad (40). A este respecto, varios de los estudios de validación analizados en este informe no han contado con estos grupos vulnerables, y se reconoce este hecho como una de sus limitaciones (103). Los criterios de exclusión que se indican con más frecuencia fueron aquellos que podían afectar a la capacidad de comprensión del paciente: el hecho de padecer alguna enfermedad mental grave (93, 95), el deterioro cognitivo o de la capacidad funcional del paciente (91, 95, 100, 102, 103), el deterioro del estado general del paciente (101) o una discapacidad intelectual (91). En otros casos, la edad (mayores de 70 años) (91) o el bajo dominio de la lengua española (99) también fueron causas de exclusión.

## 7. Conclusiones

- Para la elección de una u otra herramienta de detección precoz del malestar emocional en personas adultas con cáncer, deben considerarse aspectos como el objetivo de la detección precoz, la longitud de la escala, la fase de la enfermedad en la que se encuentran las y los pacientes, la valoración de riesgo de sesgos de los estudios y la valoración global de la escala.
- Considerando todo ello en su conjunto, destaca el Termómetro de distrés (con un punto de corte  $\geq 5$ ) como herramienta de detección precoz del malestar emocional por su brevedad (1 ítem), por haber sido validado en estudios con riesgo de sesgo bajo y cuyas muestras provienen de población oncológica en diferentes estadios de la enfermedad, así como por contar con una valoración global de la escala moderada. Esta medida de un solo ítem cuenta con una alta sensibilidad y una baja especificidad, por lo que podría considerarse un primer estadio de cribado útil para descartar distrés clínico, y combinarse posteriormente con una escala de mayor longitud, lo que contribuiría a que la especificidad del cribado en su conjunto mejorase. Para este fin, destacaría la escala HADS.
- Cabe subrayar la relevancia de emplear las herramientas de detección precoz del malestar emocional en el marco de un programa de cribado estructurado y coordinado por un equipo multidisciplinar que cuente con la figura de un o una profesional de la psicología, así como un protocolo de derivación que asegure la atención psicosocial especializada a las y los pacientes que la necesiten.
- Sería recomendable poner en marcha estudios que evalúen de manera completa la calidad psicométrica y la precisión diagnóstica de las herramientas de detección precoz del malestar emocional, ya que el número de estudios es reducido y, en algunos casos, resultan incompletos al no incluir prueba de referencia, siendo particularmente necesario en el caso del BSI-18. Asimismo, sería preciso que futuras investigaciones presten mayor atención a las cuestiones relacionadas con la implementación y la aceptabilidad de las herramientas por parte de los y las pacientes, sobre todo teniendo en cuenta la diversidad social y cultural, y las peculiaridades de los grupos desfavorecidos o de la población con discapacidad.

# Contribuciones

## Contribución de las autoras

**Patricia Gómez Salgado.** Técnica. Unidad de Asesoramiento Científico-técnico, Avalia-t; Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS – Coordinación, planificación, diseño, selección de estudios, extracción, análisis y síntesis de datos, lectura crítica de los estudios incluidos y redacción del informe.

**Yolanda Triñanes Pego.** Técnica. Unidad de Asesoramiento Científico-técnico, Avalia-t; Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS – Planificación, diseño, selección de estudios, revisión de la extracción, análisis y síntesis de datos, lectura crítica de los estudios incluidos y revisión del informe.

**Beatriz Casal Acción.** Documentalista. Unidad de Asesoramiento Científico-técnico, Avalia-t; Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS – Desarrollo de las estrategias de búsqueda, consulta en bases de datos, redacción parcial del método y revisión del informe.

**María José Faraldo Vallés.** Jefa de Servicio. Unidad de Asesoramiento Científico-técnico, Avaliat; Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS – Planificación, diseño, asesoramiento y supervisión durante todo el estudio, y revisión del informe.

## Revisión externa

**Ana María Hernández Baeza.** Profesora Titular. Metodología de las Ciencias del Comportamiento e Instituto de Investigación en Psicología de los RRHH, del Desarrollo Organizacional y de la Calidad de Vida Laboral (IDOCAL) – Universitat de València. Coordinadora de la Comisión de Tests del Consejo General de la Psicología de España (COP). En representación de la Asociación Española de Metodología de Ciencias del Comportamiento (AEMCCO).

**José Luis Herrero Burgos.** Facultativo Especialista de Área en Oncología Médica. Hospital Universitario Virgen Macarena. Servicio Andaluz de Salud. En representación de la Sociedad Española de Oncología Médica (SEOM).

**Urbano Lorenzo Seva.** Catedrático de Universidad. Metodología de las Ciencias del Comportamiento - Universitat Rovira i Virgili. En representación de la Asociación Española de Metodología de Ciencias del Comportamiento (AEMCCO).

**Paula Martínez López.** Doctora en Psicología. Psicóloga experta en Psicooncología y/o Psicología en Cuidados Paliativos. Personal Docente Investigador de la Facultad de Ciencias de la Salud de la Universidad Internacional de Valencia.

**Rebeca Mondéjar.** Facultativo Especialista de Área en Oncología Médica. Hospital Universitario de La Princesa. Servicio Madrileño de Salud. En representación de la Sociedad Española de Oncología Médica (SEOM).

**Carmen Yélamos Agua.** Psicóloga especialista en Psicología Clínica. Experta en Psicooncología y cuidados paliativos. Responsable Nacional de Atención Psicológica. En representación de la Asociación Española Contra el Cáncer.

Un **psicólogo** y dos **pacientes** de la Asociación Española Contra el Cáncer. En representación de los y las pacientes.

## Secretaría

**Noemí Raña Villar.** Administrativa. Unidad de Asesoramiento Científico-técnico, Avaliat; Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS.

**Cristina Rey Varela.** Auxiliar administrativa. Unidad de Asesoramiento Científico-técnico, Avaliat; Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS.

## Agradecimientos

Se agradece a todos los revisores y las revisoras del informe su tiempo, dedicación y comentarios que, sin duda, han redundado en la calidad del documento.

## Declaración de intereses

Las autoras y los revisores y las revisoras de este documento declaran que no ha existido ningún tipo de conflicto de interés en su realización.

# Referencias bibliográficas

1. Grassi L. Psychiatric and psychosocial implications in cancer care: the agenda of psycho-oncology. *Epidemiol Psychiatr Sci.* 2020;29:e89. PubMed PMID: 31915101.
2. Echarte Ventura A, Fuster Martí E, Roda F, Velasco Trujillo A. Estudio sobre la atención psicológica en cáncer en el Sistema Nacional de Salud en España. Madrid: Asociación Española Contra el Cáncer; 2019. Disponible en: <https://observatorio.contraelcancer.es/sites/default/files/informes/Atencion-psicologica-SNS-Espa%C3%B1a.pdf>.
3. Mitchell AJ. Short screening tools for cancer-related distress: a review and diagnostic validity meta-analysis. *J Natl Compr Canc Netw.* 2010;8(4):487-94. PubMed PMID: 20410338.
4. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines): Distress Management. Version 1.2023. NCCN; 2022. [consultado 15 dic 2022]. Disponible en: [https://www.nccn.org/professionals/physician\\_gls/pdf/distress.pdf](https://www.nccn.org/professionals/physician_gls/pdf/distress.pdf).
5. Carlson LE, Zelinski EL, Toivonen KI, Sundstrom L, Jobin CT, Damaskos P, et al. Prevalence of psychosocial distress in cancer patients across 55 North American cancer centers. *J Psychosoc Oncol.* 2019;37(1):5-21. PubMed PMID: 30592249.
6. Mehnert A, Hartung TJ, Friedrich M, Vehling S, Brahler E, Harter M, et al. One in two cancer patients is significantly distressed: Prevalence and indicators of distress. *Psycho-oncology.* 2018;27(1):75-82. PubMed PMID: 28568377.
7. Martínez López P, Andreu Vaillo Y, Galdón Garrido MJ, Romero Retes R, García-Conde Benet A, Llombart Fuertes P. Distrés emocional y problemas asociados en población oncológica adulta. *Psicooncología.* 2017;14:217-28.
8. Gao W, Bennett MI, Stark D, Murray S, Higginson IJ. Psychological distress in cancer from survivorship to end of life care: prevalence, associated factors and clinical implications. *Eur J Cancer.* 2010;46(11):203644. PubMed PMID: 20447824.
9. Carlson LE, Angen M, Cullum J, Goodey E, Koopmans J, Lamont L, et al. High levels of untreated distress and fatigue in cancer patients. *Br J Cancer.* 2004;90(12):2297-304. PubMed PMID: 15162149.
10. Zabora J, BrintzenhofeSzoc K, Curbow B, Hooker C, Piantadosi S. The prevalence of psychological distress by cancer site. *Psycho-oncology.* 2001;10(1):19-28. PubMed PMID: 11180574.
11. Estrategia en Cáncer del Sistema Nacional de Salud. Actualización aprobada por el Consejo Interterritorial del Sistema Nacional de Salud, el 24 de febrero de 2021. Madrid: Ministerio de Sanidad; 2021. [consultado 14 oct 2021]. Disponible en: <https://cpage.mpr.gob.es/producto/estrategia-en-cancer-del-sistema-nacional-de-salud-3/>.

12. Plan Europeo de Lucha contra el Cáncer. Bruselas: Comisión europea; 2021. [consultado 14 oct 2021]. Disponible en: <https://ec.europa.eu/health/sites/default/files/non-communicable-diseases/docs/eu-cancer-plan-es.pdf>.
13. NCCN practice guidelines for the management of psychosocial distress. National Comprehensive Cancer Network. Oncology (Williston Park, NY). 1999;13(5a):113-47. PubMed PMID: 10370925.
14. Howell D, Keshavarz H, Esplen MJ, Hack T, Hamel M, Howes J, et al. Pan-Canadian Practice Guideline: Screening, Assessment and Management of Psychosocial Distress, Depression and Anxiety in Adults with Cancer. Toronto;2015. [consultado 17 nov 2021]. Disponible en: <https://www.capo.ca/resources/Documents/Guidelines/3APAN--1.PDF>.
15. Syrjala KL, Jensen MP, Mendoza ME, Yi JC, Fisher HM, Keefe FJ. Psychological and behavioral approaches to cancer pain management. J Clin Oncol. 2014;32(16):1703-11. PubMed PMID: 24799497.
16. Helgeson VS, Snyder P, Seltman H. Psychological and physical adjustment to breast cancer over 4 years: identifying distinct trajectories of change. Health Psychol. 2004;23(1):3-15. PubMed PMID: 14756598.
17. Giese-Davis J, Collie K, Rancourt KM, Neri E, Kraemer HC, Spiegel D. Decrease in depression symptoms is associated with longer survival in patients with metastatic breast cancer: a secondary analysis. J Clin Oncol. 2011;29(4):413-20. PubMed PMID: 21149651.
18. Hjerl K, Andersen EW, Keiding N, Mouridsen HT, Mortensen PB, Jørgensen T. Depression as a prognostic factor for breast cancer mortality. Psychosomatics. 2003;44(1):24-30. PubMed PMID: 12515834.
19. Pinquart M, Duberstein PR. Depression and cancer mortality: a meta-analysis. Psychological medicine. 2010;40(11):1797-810. PubMed PMID: 20085667.
20. Trudel-Fitzgerald C, Savard J, Ivers H. Which symptoms come first? Exploration of temporal relationships between cancer-related symptoms over an 18-month period. Ann Behav Med. 2013;45(3):329-37. PubMed PMID: 23508468.
21. Thornton LM, Andersen BL, Crespin TR, Carson WE. Individual trajectories in stress covary with immunity during recovery from cancer diagnosis and treatments. Brain Behav Immun. 2007;21(2):185-94. PubMed PMID: 16908118.
22. Traeger L, Greer JA, Fernandez-Robles C, Temel JS, Pirl WF. Evidence-based treatment of anxiety in patients with cancer. J Clin Oncol. 2012;30(11):1197-205. PubMed PMID: 22412135.
23. Wu SM, Andersen BL. Stress generation over the course of breast cancer survivorship. J Behav Med. 2010;33(3):250-7. PubMed PMID: 20204490.

24. Lin C, Clark R, Tu P, Bosworth HB, Zullig LL. Breast cancer oral anti-cancer medication adherence: a systematic review of psychosocial motivators and barriers. *Breast Cancer Res Treat.* 2017;165(2):247-60. PubMed PMID: 28573448.
25. Mausbach BT, Schwab RB, Irwin SA. Depression as a predictor of adherence to adjuvant endocrine therapy (AET) in women with breast cancer: a systematic review and meta-analysis. *Breast Cancer Res Treat.* 2015;152(2):239-46. PubMed PMID: 26077640.
26. Yang HC, Brothers BM, Andersen BL. Stress and quality of life in breast cancer recurrence: moderation or mediation of coping? *Ann Behav Med.* 2008;35(2):188-97. PubMed PMID: 18347897.
27. Batty GD, Russ TC, Stamatakis E, Kivimäki M. Psychological distress in relation to site specific cancer mortality: pooling of unpublished data from 16 prospective cohort studies. *BMJ (Clinical research ed).* 2017;356:j108. PubMed PMID: 28122812.
28. Kissane D. Beyond the psychotherapy and survival debate: the challenge of social disparity, depression and treatment adherence in psychosocial cancer care. *Psycho-oncology.* 2009;18(1):1-5. PubMed PMID: 19097139.
29. Massie MJ. Prevalence of depression in patients with cancer. *J Natl Cancer Inst Monogr.* 2004 (32):5771. PubMed PMID: 15263042.
30. Pirl WF, Greer JA, Traeger L, Jackson V, Lennes IT, Gallagher ER, et al. Depression and survival in metastatic non-small-cell lung cancer: effects of early palliative care. *J Clin Oncol.* 2012;30(12):1310-5. PubMed PMID: 22430269.
31. Andersen BL, DeRubeis RJ, Berman BS, Gruman J, Champion VL, Massie MJ, et al. Screening, assessment, and care of anxiety and depressive symptoms in adults with cancer: an American Society of Clinical Oncology guideline adaptation. *J Clin Oncol.* 2014;32(15):1605-19. PubMed PMID: 24733793.
32. Krebber AM, Jansen F, Cuijpers P, Leemans CR, Verdonck-de Leeuw IM. Screening for psychological distress in follow-up care to identify head and neck cancer patients with untreated distress. *Support Care Cancer.* 2016;24(6):2541-8. PubMed PMID: 26694718.
33. Funk R, Cisneros C, Williams RC, Kendall J, Hamann HA. What happens after distress screening? Patterns of supportive care service utilization among oncology patients identified through a systematic screening protocol. *Support Care Cancer.* 2016;24(7):2861-8. PubMed PMID: 26838023.
34. Hernández M, Cruzado JA. La atención psicológica a pacientes con cáncer: de la evaluación al tratamiento. *Clínica y Salud.* 2013;24:1-9.

35. Andreu Vaíllo Y, Galdón Garrido MJ, Martínez López P, Martínez Arroyo O. Precisión diagnóstica del termómetro de distrés en neoplasias hematológicas. *Psicooncología*. 2015;12(2-3):237-47.
36. European Guide for Quality National Cancer Control Programmes. Ljubljana: National Institute of Public Health; 2015. Disponible en: <https://cancercontrol.eu/archived/uploads/images/European-Guide-for-Quality-National-Cancer-Control-Programmes-web.pdf>.
37. Holland J, Watson M, Dunn J. The IPOS new International Standard of Quality Cancer Care: integrating the psychosocial domain into routine care. *Psycho-oncology*. 2011;20(7):677-80. PubMed PMID: 21710503.
38. European Guide on Quality Improvement in Comprehensive Cancer Control. Brussels, Belgium: National Institute of Public Health; 2017 [consultado 13 oct 2021]. Disponible en: <https://cancercontrol.eu/archived/uploads/images/Guide/pdf/CanCon-Guide-FINAL-Web.pdf>.
39. A Dictionary of Epidemiology. 5<sup>th</sup> ed. Oxford: Oxford University Press; 2014. Early detection of disease. [consultado 06 jun 2022]. Disponible en: <https://www.oxfordreference.com/view/10.1093/acref/9780195314496.001.0001/acref-9780195314496-e-579?rskey=zUnWUO&result=579>
40. Grupo de trabajo de la Ponencia de Cribado de la Comisión de Salud Pública. Documento marco sobre cribado poblacional. Madrid: Ministerio de Sanidad y Política Social; 2010 [consultado 25 nov 2021]. Disponible en: <https://www.msbs.gob.es/en/profesionales/saludPublica/prevPromocion/Cribado/docs/Cribado-poblacional.pdf>.
41. García-Campayo J, Pérez-Yus MC, García-Bustinduy M, Daudén E. Early Detection of Emotional and Behavioral Disorders in Dermatology. *Actas Dermosifiliogr*. 2016;107(4):294-300. PubMed PMID: 26651326.
42. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Estándares para pruebas educativas y psicológicas. Washington: American Educational Research Association; 2018.
43. Muzzatti B, Annunziata MA. Psychological distress screening in cancer patients: psychometric properties of tools available in Italy. *Tumori*. 2012;98(4):501-9. PubMed PMID: 23052168.
44. Mitchell AJ. Pooled results from 38 analyses of the accuracy of distress thermometer and other ultra-short methods of detecting cancer-related mood disorders. *J Clin Oncol*. 2007;25(29):4670-81. PubMed PMID: 17846453.
45. Shinn EH, Valentine A, Baum G, Carmack C, Kilgore K, Bodurka D, et al. Comparison of four brief depression screening instruments in ovarian cancer patients: Diagnostic accuracy using traditional versus alternative cutpoints. *Gynecologic oncology*. 2017;145(3):562-8. PubMed PMID: 28400146.

46. Syrjala KL, Sutton SK, Jim HS, Knight JM, Wood WA, Lee SJ, et al. Cancer and treatment distress psychometric evaluation over time: A BMT CTN 0902 secondary analysis. *Cancer*. 2017;123(8):1416-23. PubMed PMID: 27893933.
47. Alberta Health Services. Cancer Pain. Alberta: Alberta Health Services; 2018. Informe N.º: SUPP-012. [consultado 16 nov 2021]. Disponible en: <https://www.albertahealthservices.ca/assets/info/hp/cancer/if-hp-cancer-guide-supp012-cancer-pain.pdf>
48. Vodermaier A, Linden W, Siu C. Screening for emotional distress in cancer patients: a systematic review of assessment instruments. *Journal of the National Cancer Institute*. 2009;101(21):1464-88. PubMed PMID: 19826136.
49. Caminiti C, Diodati F, Annunziata MA, Di Giulio P, Isa L, Mosconi P, et al. Psychosocial Care for Adult Cancer Patients: Guidelines of the Italian Medical Oncology Association. *Cancers (Basel)*. 2021;13(19). PubMed PMID: 34638360.
50. Fallon M, Giusti R, Aielli F, Hoskin P, Rolke R, Sharma M, et al. Management of cancer pain in adult patients: ESMO Clinical Practice Guidelines. *Ann Oncol*. 2018;29(Suppl 4):iv166-iv91. PubMed PMID: 30052758.
51. National Institute for Clinical Excellence (NICE). Improving supportive and palliative care for adults with cancer. Londres: National Institute for Clinical Excellence (NICE); 2004. Disponible en: <https://www.nice.org.uk/guidance/csg4/resources/improving-supportive-and-palliative-care-for-adults-with-cancer-pdf-773375005>
52. Roth AJ, Kornblith AB, Batel-Copel L, Peabody E, Scher HI, Holland JC. Rapid screening for psychologic distress in men with prostate carcinoma: a pilot study. *Cancer*. 1998;82(10):1904-8. PubMed PMID: 9587123.
53. Ma X, Zhang J, Zhong W, Shu C, Wang F, Wen J, et al. The diagnostic role of a short screening tool--the distress thermometer: a meta-analysis. *Support Care Cancer*. 2014;22(7):1741-55. PubMed PMID: 24510195.
54. Bruera E, Kuehn N, Miller MJ, Selmsler P, Macmillan K. The Edmonton Symptom Assessment System (ESAS): a simple method for the assessment of palliative care patients. *J Palliat Care*. 1991;7(2):6-9. PubMed PMID: 1714502.
55. Watanabe SM, Nekolaichuk CL, Beaumont C. The Edmonton Symptom Assessment System, a proposed tool for distress screening in cancer patients: development and refinement. *Psycho-oncology*. 2012;21(9):977-85. PubMed PMID: 21671304.
56. Nekolaichuk C, Beaumont C, Watanabe S, Buttenschoen D. Edmonton Symptom Assessment System – Revised (ESAS-r) Administration Manual: Alberta Health Services (AHS); 2019.
57. Selby D, Cascella A, Gardiner K, Do R, Moravan V, Myers J, et al. A single set of numerical cutpoints to define moderate and severe symptoms for the Edmonton

- Symptom Assessment System. *J Pain Symptom Manage.* 2010;39(2):241-9. PubMed PMID: 19963335.
58. Hui D, Bruera E. The Edmonton Symptom Assessment System 25 Years Later: Past, Present, and Future Developments. *J Pain Symptom Manage.* 2017;53(3):630-43. PubMed PMID: 28042071.
  59. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care.* 2003;41(11):1284-92. PubMed PMID: 14583691.
  60. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606-13. PubMed PMID: 11556941.
  61. Manea L, Gilbody S, Hewitt C, North A, Plummer F, Richardson R, et al. Identifying depression with the PHQ-2: A diagnostic meta-analysis. *J Affect Disord.* 2016;203:382-95. PubMed PMID: 27371907.
  62. The Screening for Distress initiative of the Cancer Journey Action Group. Moving towards person-centered care. Part A: Background, Recommendations, and Implementation. *Guide to Implementing Screening for Distress, the 6<sup>th</sup> Vital Sign.* Toronto: Canadian Partnership Against Cancer; 2009.
  63. The Screening for Distress initiative of the Cancer Journey Portfolio. Screening for distress, the 6<sup>th</sup> vital sign: A guide to implementing best practices in person-centred care. Toronto: Canadian Partnership Against Cancer; 2012. [consultado 07 dic 2021]. Disponible en: <https://www.virtualhospice.ca/Assets/Distress-%20CPAC-20150713161546.pdf>
  64. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006;166(10):1092-7. PubMed PMID: 16717171.
  65. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV).* Washington (DC): American Psychiatric Association; 1994.
  66. Instruction Manual. Instructions for Patient Health Questionnaire (PHQ) and GAD-7 Measures [Internet]. [consultado 02 abr 2022]. Disponible en: <https://www.phqscreeners.com/images/sites/g/files/g10016261/f/201412/instructions.pdf>
  67. Kroenke K, Spitzer RL, Williams JB, Löwe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *Gen Hosp Psychiatry.* 2010;32(4):345-59. PubMed PMID: 20633738.
  68. Williams N. The GAD-7 questionnaire. *Occupational Medicine.* 2014;64(3):224.
  69. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD2: a systematic review and diagnostic metaanalysis. *Gen Hosp Psychiatry.* 2016;39:24-31.

70. Chamberlain JM. Which prescriptive screening programmes are worth while? *J Epidemiol Community Health*. 1984;38(4):270-7. PubMed PMID: 6239897.
71. WHO Regional Office for Europe. Screening programmes: a short guide. Increase effectiveness, maximize benefits and minimize harm. Copenhagen; 2020. [consultado 10 dic 2021]. Disponible en: <https://apps.who.int/iris/bitstream/handle/10665/330829/9789289054782-eng.pdf>
72. Bultz BD, Carlson LE. Emotional distress: the sixth vital sign in cancer care. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2005;23(26):6440-1. PubMed PMID: 16155033.
73. Bultz BD, Groff SL, Fitch M, Blais MC, Howes J, Levy K, et al. Implementing screening for distress, the 6th vital sign: a Canadian strategy for changing practice. *Psycho-oncology*. 2011;20(5):463-9. PubMed PMID: 21456060.
74. Dolbeault S, Boistard B, Meuric J, Copel L, Bredart A. Screening for distress and supportive care needs during the initial phase of the care process: A qualitative description of a clinical pilot experiment in a French cancer center. *Psycho-oncology*. 2011;20(6):585-93.
75. Grassi L, Rossi E, Caruso R, Nanni MG, Pedrazzi S, Sofritti S, et al. Educational intervention in cancer outpatient clinics on routine screening for emotional distress: an observational study. *Psycho-oncology*. 2011;20(6):669-74. PubMed PMID: 21370316.
76. Okuyama T, Kizawa Y, Morita T, Kinoshita H, Uchida M, Shimada A, et al. Current Status of Distress Screening in Designated Cancer Hospitals: A Cross-Sectional Nationwide Survey in Japan. *Journal of the National Comprehensive Cancer Network : JNCCN*. 2016;14(9):1098-104. PubMed PMID: 27587622.
77. van Nuenen FM, Donofrio SM, Tuinman MA, van de Wiel HB, Hoekstra-Weebers JE. Feasibility of implementing the 'Screening for Distress and Referral Need' process in 23 Dutch hospitals. *Support Care Cancer*. 2017;25(1):103-10. PubMed PMID: 27565789.
78. Zebrack B, Kayser K, Bybee D, Padgett L, Sundstrom L, Jobin C, et al. A Practice-Based Evaluation of Distress Screening Protocol Adherence and Medical Service Utilization. *Journal of the National Comprehensive Cancer Network : JNCCN*. 2017;15(7):903-12. PubMed PMID: 28687578.
79. Institute of Medicine, Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. *Clinical Practice Guidelines We Can Trust*. Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E, editors. Washington (DC): National Academies Press (US); 2011.
80. López de Argumedo M, Reviriego E, Gutiérrez A, Bayón JC. Actualización del Sistema de Trabajo Compartido para Revisiones Sistemáticas de la Evidencia

Científica y Lectura Crítica (Plataforma FLC 3.0). Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad 2017.

81. Campo-Arias A, Oviedo HC. Propiedades psicométricas de una escala: la consistencia interna. *Revista de Salud Pública*. 2008;10(5):831-9.
82. Henson RK. Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development*. 2001;34(3):177-89.
83. Arias A, Sireci S. Validez y Validación para Pruebas Educativas y Psicológicas. *Revista Iberoamericana de Psicología*. 2021;14(1).
84. Elosua P. Sobre la validez de los tests. *Psicothema*. 2003;15(2):315-21.
85. Carvajal A, Centeno C, Watson R, Martínez M, Sanz Rubiales Á. ¿Cómo validar un instrumento de medida de la salud? *Anales del Sistema Sanitario de Navarra*. 2011;34:63-72.
86. Pérez Gil JA, Chacón Moscoso S, Moreno Rodríguez R. Validez de constructo: el uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema*. 2000;12:442-6.
87. Puñal Riobóo J, Baños Álvarez E, Varela Lema L, Castillo Muñoz MA, Atienza Merino G, Ubago Pérez R, et al. Guía para la elaboración y adaptación de informes rápidos de evaluación de tecnologías sanitarias. Santiago de Compostela: Agencia Gallega para la Gestión del Conocimiento en Salud, ACIS; 2016. Informe N.º: avalia-t 2015/10. Disponible en: <http://hdl.handle.net/20.500.11940/9007>
88. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine*. 2011;155(8):52936. PubMed PMID: 22007046.
89. Ciapponi A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. *Evidencia Actualización en la Práctica Ambulatoria*. 2015;18(1):22-30.
90. Galdón MJ, Durá E, Andreu Y, Ferrando M, Murgui S, Pérez S, et al. Psychometric properties of the Brief Symptom Inventory-18 in a Spanish breast cancer sample. *Journal of psychosomatic research*. 2008;65(6):533-9. PubMed PMID: 19027441.
91. Priede A, Ruiz Torres M, Hoyuela F, Herrán A, González-Blanch C. El termómetro del distrés como prueba de cribado de ansiedad y depresión en pacientes oncológicos recién diagnosticados. *Psicooncología*. 2014;11(1):31-43.
92. Martínez López P, Conchado Peiró A, Andreu Vaillo Y, Galdón Garrido MJ. Psychometric properties of the Brief Symptom Inventory-18 in a heterogeneous sample of adult cancer patients. *Revista Latinoamericana de Psicología*. 2019;51(1):1-8.

93. Calderon C, Ferrando PJ, Lorenzo-Seva U, Hernández R, Oporto-Alonso M, Jiménez-Fonseca P. Factor structure and measurement invariance of the Brief Symptom Inventory (BSI-18) in cancer patients. *Int J Clin Health Psychol.* 2020;20(1):71-80.
94. López-Roig S, Terol MC, Pastor MA, Neipp MC, Massutí B, Rodríguez-Marín J, et al. Ansiedad y Depresión. Validación de la escala HAD en pacientes oncológicos. *Revista de Psicología de la Salud.* 2000;12(2):127-55.
95. Costa G, Pérez X, Salamero M, Gil FL. Discriminación del malestar emocional en pacientes oncológicos utilizando la escala de Ansiedad y Depresión hospitalaria (HADS). *Ansiedad y Estrés.* 2009;15(2-3):217-29.
96. Martínez P, Durá E, Andreu Y, Galdón MJ, Murgui S, Ibáñez E. Structural validity and distress screening potential of the Hospital Anxiety and Depression Scale in cancer. *International Journal of Clinical and Health Psychology.* 2012;12(3):435-47.
97. Martínez P, Galdón MJ, Andreu Y, Ibáñez E. The Distress Thermometer in Spanish cancer patients: convergent validity and diagnostic accuracy. *Support Care Cancer.* 2013;21(11):3095-102. PubMed PMID: 23832312.
98. Martínez P, Andreu Y, Galdón MJ, Ibáñez E. Improving the diagnostic accuracy of the Distress Thermometer: A potential role for the Impact Thermometer. *J Pain Symptom Manage.* 2015;50(1):124-9.
99. Andreu Vaillo Y, Martínez Lopez P, Galdón Garrido MJ. Use of the Distress Thermometer in Cancer Survivors: Convergent Validity and Diagnostic Accuracy in a Spanish Sample. *Oncology nursing forum.* 2019;46(4):442-50.
100. Carvajal A, Hribernik N, Duarte E, Sanz-Rubiales A, Centeno C. The Spanish version of the Edmonton Symptom Assessment System-revised (ESAS-r): First psychometric analysis involving patients with advanced cancer. *J Pain Symptom Manage.* 2013;45(1):129-36.
101. Martínez García M, Porta Sales J, Espinosa Rojas J, Gómez Batiste-Alentorn X. Evaluación sencilla y rápida del estrés emocional en una consulta externa de Cuidados Paliativos. *Medicina Paliativa.* 2006;13:144-8.
102. Limonero JT, Mateo D, Maté-Mendez J, González-Barboteo J, Bayés R, Bernaus M, et al. Evaluación de las propiedades psicométricas del cuestionario de Detección de Malestar Emocional (DME) en pacientes oncológicos. *Gac Sanit.* 2012;26(2):145-52.
103. Mateo-Ortega D, Limonero JT, Maté-Méndez J, Beas E, González-Barboteo J, Barbero E, et al. Development of a tool to identify and assess psychosocial and spiritual needs in end-of-life patients: The ENP-E scale. *Palliat Support Care.* 2019;17(4):441-7. PubMed PMID: 30322421.
104. Derogatis LR. Brief Symptom Inventory 18. Administration, scoring and procedures manual. Minneapolis (MN): National Computer Systems Pearson; 2000.

105. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*. 1983;67:361-70.
106. Akizuki N, Yamawaki S, Akechi T, Nakano T, Uchitomi Y. Development of an Impact Thermometer for use in combination with the Distress Thermometer as a brief screening tool for adjustment disorders and/or major depression in cancer patients. *J Pain Symptom Manage*. 2005;29(1):91-9. PubMed PMID: 15652442.
107. Maté Méndez J, Mateo Ortega D, Bayés R, Bernaus Martí M, Casas C, González Barboteo J, et al. Elaboración y propuesta de un instrumento para la detección de malestar emocional en enfermos al final de la vida. *Psicooncología*. 2009;6(2-3):507-18.
108. Spielberger CD, Gorsuch RL, Lushene RE. *Manual for the State-Trait Anxiety Inventory (Self-Evaluation Questionnaire)*. Palo Alto (CA): Consulting Psychologists Press (Versión Castellana TEA Ediciones, 1982) 1970.
109. Conde V, Escribá J, Izquierdo J. Evaluación Estadística y Adaptación Castellana de la Escala Autoaplicada para la Depresión, de Zung. *Archives of Neurobiology*. 1970;23(2):185-206.
110. First MB, Spitzer RL, Gibbon M, Williams JBW. *Entrevista Clínica Estructurada para los Trastornos del Eje I del DSM-IV. SCID-I*. Barcelona: Masson; 1999.
111. Herrero MJ, Blanch J, Peri JM, De Pablo J, Pintor L, Bulbena A. A validation study of the hospital anxiety and depression scale (HADS) in a Spanish population. *Gen Hosp Psychiatry*. 2003;25:277-83.
112. Tejero A, Guimerá E, Farré J, Peri J. Uso clínico del HADS (Hospital Anxiety and Depression Scale) en población psiquiátrica: un estudio de su sensibilidad, fiabilidad y validez. *Revista del Departamento de Psiquiatría de la Facultad de Medicina de Barcelona*. 1986;13(5):233-8.
113. Groenvold M, Petersen MA, Aaronson NK, Arraras JI, Blazeby JM, Bottomley A, et al. The development of the EORTC QLQ-C15-PAL: a shortened questionnaire for cancer patients in palliative care. *Eur J Cancer*. 2006;42(1):55-64. PubMed PMID: 16162404.
114. Bayes R, Limonero JT, Barreto P, Comas MD. Assessing suffering. *Lancet*. 1995;346(8988):1492. PubMed PMID: 7491014.
115. Bayés R, Limonero JT, Barreto P, Comas MD. A way to screen for suffering in palliative care. *J Palliat Care*. 1997;13(2):22-6. PubMed PMID: 9231584.
116. Oviedo HC, Campo-Arias A. Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría*. 2005;34:572-80.
117. Kaiser HF, Rice J. Little Jiffy, Mark Iv. *Educational and Psychological Measurement*. 1974;34(1):111-7.

118. Wheaton B, Muthén B, Alwin DF, Summers GF. Assessing Reliability and Stability in Panel Models. *Sociological Methodology*. 1977;8:84-136.
119. Brooke PP, Russell DW, Price JL. Discriminant validation of measures of job satisfaction, job involvement, and organizational commitment. *Journal of Applied Psychology*. 1988 (73):139-45.
120. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*. 1999;6(1):1-55.
121. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*. 1980;88(3):588-606.
122. Byrne B. *Structural equation modeling with AMOS: Basic concepts, applications and programming*. Mahwah, NJ: Erlbaum; 2001.
123. Hinkle DE, Wiersma W, Jurs SG. *Applied Statistics for the Behavioral Sciences*. 5th ed. Boston: Houghton Mifflin; 2003.
124. Vizcaíno-Salazar GJ. Importancia del cálculo de la sensibilidad, la especificidad y otros parámetros estadísticos en el uso de las pruebas de diagnóstico clínico y de laboratorio. *Medicina y Laboratorio*. 2017;23(7-8):365-86.
125. Florkowski CM. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *Clin Biochem Rev*. 2008;29 Suppl 1(Suppl 1):S83-S7. PubMed PMID: 18852864.
126. Li F, He H. Assessing the Accuracy of Diagnostic Tests. *Shanghai Arch Psychiatry*. 2018;30(3):207-12. PubMed PMID: 30858674.
127. Hernández A, Elosua P, Fernández-Hermida JR, Muñoz J. Comisión de Test: veinticinco años velando por la calidad de los test. *Papeles del Psicólogo*. 2022;43:55-62.
128. Mitchell AJ. Screening for cancer-related distress: When is implementation successful and when is it unsuccessful? *Acta Oncol*. 2013;52(2):216-24. PubMed PMID: 23320770.
129. García Aguado J. Cribar o no cribar: ¿esa es la cuestión? En: Asociación Española de Pediatría de Atención Primaria, editor. *Curso de Actualización Pediatría 2017*. Madrid: Lúa Ediciones 3.0; 2017. p. 21121.
130. Martínez López P. *Cribado de distrés emocional en población oncológica adulta: validación de instrumentos [Tesis Doctoral]*. Valencia: Universidad de Valencia; 2016.
131. Muñoz J, Fonseca-Pedrero E. Diez pasos para la construcción de un test. *Psicothema*. 2019;31(1):716. PubMed PMID: 30664405.
132. Hausner E, Waffenschmidt S, Hafstad E, Harboe I, Isabel-Gómez R. *Process of information retrieval for systematic reviews and health technology assessments on clinical effectiveness*. Diemen (The Netherlands): EUnetHTA; 2019.

133. Terol-Cantero MC, Cabrera-Perona V, Martín-Aragón M. Revisión de estudios de la Escala de Ansiedad y Depresión Hospitalaria (HAD) en muestras españolas. *Anales de Psicología*. 2015;31(2):494-503.
134. Clover K, Carter GL, Mackinnon A, Adams C. Is my patient suffering clinically significant emotional distress? Demonstration of a probabilities approach to evaluating algorithms for screening for distress. *Support Care Cancer*. 2009;17(12):1455-62. PubMed PMID: 19277724.
135. Singer S, Kuhnt S, Götze H, Hauss J, Hinz A, Liebmann A, et al. Hospital anxiety and depression scale cutoff scores for cancer patients in acute care. *Br J Cancer*. 2009;100(6):908-12. PubMed PMID: 19240713.
136. Herschbach P, Britzelmeir I, Dinkel A, Giesler JM, Herkommer K, Nest A, et al. Distress in cancer patients: Who are the main groups at risk? *Psycho-oncology*. 2020;29(4):703-10.
137. Grupo de trabajo de la Guía de Práctica Clínica en Cáncer Hereditario. Guía de Práctica Clínica en Cáncer Hereditario de la Comunitat Valenciana. 3ª edición. Servicio de Promoción de la Salud y Prevención en Entorno Sanitario - Conselleria de Sanitat Universal i Salut Pública - Generalitat Valenciana; 2017. [consultado 19 dic 2022]. Disponible en: <https://portal.guiasalud.es/wp-content/uploads/2018/12/GPC-563-Cancer-Hereditario-Comunitat-Valenciana-2017.pdf>
138. Sociedad Oncológica de Galicia. I Guía de Cáncer Hereditario de Galicia. Madrid: Luzán 5, S.A.; 2011.
139. van Scheppingen C, Schroevers MJ, Pool G, Smink A, Mul VE, Coyne JC, et al. Is implementing screening for distress an efficient means to recruit patients to a psychological intervention trial? *Psycho-oncology*. 2014;23(5):516-23.
140. Schouten B, Avau B, Bekkering GTE, Vankrunkelsven P, Mebis J, Hellings J, et al. Systematic screening and assessment of psychosocial well-being and care needs of people with cancer. *Cochrane Database Syst Rev*. 2019;3:CD012387. PubMed PMID: 30909317.
141. Dilworth S, Higgins I, Parker V, Kelly B, Turner J. Patient and health professional's perceived barriers to the delivery of psychosocial care to adults with cancer: A systematic review. *Psycho-oncology*. 2014;23(6):601-12.

# Anexos

## Anexo A. Estrategia de búsqueda

### BÚSQUEDA ESPECÍFICA DE GUÍAS DE PRÁCTICA CLÍNICA E INFORMES DE EVALUACIÓN DE TECNOLOGÍAS SANITARIAS

La búsqueda bibliográfica para la localización de guías de práctica clínica e informes de ETS se realizó en noviembre de 2021 con una estrategia de búsqueda específica, en las siguientes bases de datos:

#### GuiaSalud

	Términos de búsqueda	N.º de ítems
1	especialidades	
2	Oncología médica	1
3	Oncología radioterápica	1

#### G-I-N

	Términos de búsqueda	N.º de ítems
1	distress cancer	1
2	cancer wellbeing	0
3	cancer well-being	0
4	oncology distress	0
5	oncology wellbeing	0
6	oncology well-being	0
7	oncology stress	0
8	cancer psychology	1
9	PROSTATE CANCER	36
10	colorectal CANCER	41
11	lung cancer 2001-	55
12	breast cancer 2001-	59

#### NICE

	Términos de búsqueda	N.º de ítems
1	emotional distress cancer	3
2	psychological distress cancer	11
3	psychosocial distress cancer	4
4	eliminados/duplicados	15

## Tripdatabase

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
#1	"emotional distress" AND oncology AND (evaluation OR test OR "early detection" OR tool OR survey OR "quality of life" OR questionnaire*)	57
#2	"psychosocial distress" AND oncology AND (evaluation OR test OR "early detection" OR tool OR survey OR "quality of life" OR questionnaire*)	49
#3	"emotional distress" cancer (tool OR survey OR "Quality of Life" OR Questionnaire*) filtering Guidelines	96
#4	"emotional distress" cancer (tool OR survey OR "Quality of Life" OR Questionnaire*) Filtering Systematic reviews	36
#5	"psychological distress" AND cancer AND (tool OR survey OR "Quality of Life" OR Questionnaire*) filtering systematic reviews	93
#6	"psychological distress" AND cancer AND ("Quality of Life") filtering by guidelines	58
#7	"emotional distress" AND oncology AND (evaluation OR test OR "early detection" OR tool OR survey OR "quality of life" OR questionnaire*)	57

## International HTA database

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
#1	"Psychological Distress"[mh] or ((emotional OR psychological or psychosocial or psycho-social) AND (distress OR stress OR wellbeing or well-being))	437
#2	questionnaire* or interview* or survey* or scale* or Psychometrics or "Quality of Life"	2009
#3	Neoplasia OR Neoplasias OR Neoplasm OR Tumors OR Tumor OR Cancer OR Cancers OR Malignancy OR Malignancies	2810
#4	#1 AND #2 AND #3	23

## Medline, Premedline (Pubmed)

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
#1	(emotion*[TIAB] OR psychologic*[TIAB]) AND (wellbeing[TIAB] or well-being[TIAB] OR distress*[TIAB] OR stress*[TIAB]) AND (Thermometer[TIAB] OR screen*[TIAB] OR assess*[TIAB] OR evaluation OR test OR "early detection")	75.795
#2	("Patient Outcome Assessment"[Mesh] OR "Interview" [Publication Type] OR "Interviews as Topic"[Mesh] OR "Interview, Psychological"[Mesh] OR "Surveys and Questionnaires"[Mesh] OR "Psychiatric Status Rating Scales"[Mesh] OR questionnaire*[TIAB] OR interview*[TIAB] OR survey*[TIAB] or scale*[TIAB]) AND (psychological*[TIAB] OR psychosocial[TIAB] OR psycho-social[TIAB]) Sort by: Most Recent	156.912
#3	Neoplasms[Mesh] OR Neoplasia[TIAB] OR Neoplasias[TIAB] OR Neoplasm*[TIAB] OR Tumor[TIAB] OR Tumour[TIAB] OR Tumors[TIAB] OR Tumours[TIAB] OR Cancer[TIAB] OR Cancers[TIAB] OR Malignancy[TIAB] OR Malignancies[TIAB] OR oncolog*[TIAB]	4.560.541
#4	#1 AND #2 AND #3 Filters applied: Consensus Development Conference, Consensus Development Conference, NIH, Guideline, Practice Guideline	4

## Embase (Ovid)

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
1	exp emotional stress/	22.470
2	((emotional or psychological) adj4 (wellbeing or well-being or distress* or stress* or evaluation or test or "early detection") adj4 (Thermometer or screen* or assess*)) .ab,hw,kw,sh,ti.	5652
3	1 or 2	27.562
4	*neoplasm/	119.931

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
5	(Neoplasia or Neoplasias or Neoplasm or Tumors or Tumor or Cancer or Cancers or Malignancy or Malignancies).ab,hw,kw,sh,ti.	4.690.591
6	4 or 5	4.690.591
7	health survey/ or health care survey/	213.116
8	*outcome assessment/	36.213
9	*interview/ or *psychological interview/	6923
10	exp psychological rating scale/	18.487
11	(questionnaire* or interview* or survey* or scale*).ab,hw,kw,ti.	3.322.979
12	(psychological* or psychosocial or psycho-social).ab,hw,kw,ti.	684.471
13	7 or 8 or 9 or 10 or 11	3.348.780
14	12 and 13	303.413
15	3 and 6 and 14	1575
16	limit 15 to (embase and (adult <18 to 64 years> or aged <65+ years>))	702
17	exp *practice guideline/	103.620
18	16 and 17	3

#### APA PsycInfo (Ovid)

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
1	(wellbeing or well-being or distress* or stress*).ab,hw,mh,ti.	413.522
2	(emotional or psychological).ab,hw,mh,ti.	549.092
3	(Thermometer or screen* or assess* or evaluation or test or "early detection").ab,hw,mh,ti.	1.257.564
4	1 and 2 and 3	58.757
5	exp psychological stress/	7550
6	4 or 5	63.424
7	exp Neoplasms/	53.600
8	(Neoplasia or Neoplasias or Neoplasm or Tumors or Tumor or Cancer or Cancers or Malignancy or Malignancies).ab,hw,mh,ti.	75.701
9	7 or 8	80.360
10	exp "Quality of Life"/ or exp Surveys/ or exp Interviews/ or exp Psychometrics/	245.619
11	((questionnaire* or interview* or survey* or scale*) and (psychological* or psychosocial or psycho-social)).ab,hw,mh,ti.	141.427
12	10 or 11	363.015
13	6 and 9 and 12	2538
14	limit 13 to ("300 adulthood <age 18 yrs and older>" or 320 young adulthood <age 18 to 29 yrs> or 340 thirties <age 30 to 39 yrs> or 360 middle age <age 40 to 64 yrs> or "380 aged <age 65 yrs and older>" or "390 very old <age 85 yrs and older>")	2077
15	exp Evidence Based Practice/	19.191
16	14 and 15	3
17	("practice guideline" or "consensus").ab,hw,ti.	30.281
18	14 and 17	2
19	16 or 18	5

## CINAHL (EBSCO)

	Términos de búsqueda	N.º de ítems
S1	TI ( emotional or psychological ) AND TI ( wellbeing or well-being or distress* or stress* ) AND TI ( Thermometer or screen* or assess* or evaluation or test or "early detection" )	448
S2	AB ( emotional or psychological ) N3 AB ( wellbeing or well-being or distress* or stress* ) N3 AB ( Thermometer or screen* or assess* or evaluation or test or "early detection" )	1123
S3	S1 OR S2	448
S4	MH "Psychological Distress"	2092
S5	S3 OR S4	2532
S6	TI ( (questionnaire* or interview* or survey* or scale*) N4 (psychological* or psychosocial or psycho-social) )	461
S7	AB ( (questionnaire* or interview* or survey* or scale*) N4 (psychological* or psychosocial or psycho-social) )	6297
S8	S6 OR S7	6561
S9	S5 AND S8	143
S10	MH "Neoplasms" OR TI ( Neoplasia OR Neoplasias OR Neoplasm OR Tumors OR Tumor OR Cancer OR Cancers OR Malignancy OR Malignancies ) OR AB ( Neoplasia OR Neoplasias OR Neoplasm OR Tumors OR Tumor OR Cancer OR Cancers OR Malignancy OR Malignancies )	591.708
S11	S9 AND S10 Especificar por SubjectAge: - all adult	11

## BÚSQUEDA ESPECÍFICA DE ESTUDIOS PRIMARIOS

La búsqueda bibliográfica para la localización de estudios primarios se realizó en diciembre de 2021 con una estrategia de búsqueda específica, en las siguientes bases de datos:

Medline, Premedline (OVID)

	Términos de búsqueda	N.º de ítems
1	((emotion* or psychologic* or psychosocial) adj5 (distress* or stress*)).ab,kw,ti.	69.385
2	Stress, Psychological/di, ep, px [Diagnosis, Epidemiology, Psychology]	43.455
3	Psychological Distress/	2696
4	1 or 2 or 3	96.306
5	Neoplasms/ep, px [Epidemiology, Psychology]	52.103
6	(Neoplasia or Neoplasias or Neoplasm* or Tumor or Tumour or Tumors or Tumours or Cancer or Cancers oncolog*).ab,kw,ti.	3.095.356
7	5 or 6	3.102.595
8	manifest anxiety scale/ or test anxiety scale/	1096
9	**surveys and questionnaires"/ or *health care surveys/ or *health surveys/ or *patient health questionnaire/ or *self report/	72.065
10	Sensitivity and Specificity/	360.673
11	Psychometrics/ or Interviews as Topic/mt, px, st or Patient Outcome Assessment/ or Interview, Psychological/	105.861
12	(test* or scale* or questionnaire* or instrument* or tool* or inventor* or Thermometer* or intervent* or surve*).ab,kw,ti.	6.719.637
13	8 or 9 or 10 or 11 or 12	6.946.382
14	Spain/	84.081
15	(spain or spanish or espa?a or espa?ol).ab,kw,ti.	103.896

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
16	14 or 15	138.172
17	(develop* or validat* or construct* or "psychometric properties" or "diagnostic accuracy" or screen* or adapt* or translat* or evaluat* or assess* or detect*).ab,kw,ti.	12.376.476
18	4 and 7 and 13 and 16 and 17	73
19	limit 18 to ("all adult (19 plus years)" or "adolescent (13 to 18 years)" or "young adult (19 to 24 years)" or "adult (19 to 44 years)" or "young adult and adult (19-24 and 19-44)" or "middle age (45 to 64 years)" or "middle aged (45 plus years)" or "all aged (65 and over)" or "aged (80 and over)")	64

### Embase (Ovid)

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
1	((emotion* or psycholog* or psychosocial) adj5 (distress* or stress*)).ab,kw,ti.	91.832
2	physiological stress/di, ep [Diagnosis, Epidemiology]	5
3	distress syndrome/di, ep [Diagnosis, Epidemiology]	2290
4	1 or 2 or 3	84.071
5	neoplasm/	429.145
6	(Neoplasia or Neoplasias or Neoplasm* or Tumor or Tumour or Tumors or Tumours or Cancer or Cancers oncolog*).ab,kw,ti.	4.136.569
7	5 or 6	4.188.464
8	psychometry/ or rating scale/	178.060
9	health survey/ or short survey/ or interview/	784.503
10	questionnaire/ or "quality of life"/	1.204.959
11	self report/	134.512
12	sensitivity and specificity/	413.046
13	(test* or scale* or questionnaire* or instrument* or tool* or inventor* or Thermometer* or intervent* or surve* or screen*).ab,kw,ti.	9.453.519
14	8 or 9 or 10 or 11 or 12 or 13	10.566.643
15	spain/ or balearic islands/ or basque country/ or canary islands/ or catalonia/ or "galicia (spain)"/ or navarre/ or valencian community/	105.034
16	(spain or spanish or espa?a or espa?ol).ab,kw,ti.	147.478
17	15 or 16	177.204
18	(develop* or validat* or construct* or "psychometric properties" or "diagnostic accuracy" or screen* or adapt* or translat* or evaluat* or assess* or detect*).ab,kw,ti.	15.979.079
19	4 and 7 and 14 and 17 and 18	122
20	limit 19 to (adult <18 to 64 years> or aged <65+ years>)	76

### APA PsycInfo (Ovid)

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
1	((emotion* or psycholog* or psychosocial) adj5 (distress* or stress*)).ab,hw,ti.	64.867
2	exp Distress/	25.204
3	(emotion* or psycholog* or psychosocial).mp.	974.378
4	2 and 3	19.568
5	1 or 4	70.293

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
6	exp Neoplasms/	56.038
7	(Neoplasia or Neoplasias or Neoplasm* or Tumor or Tumour or Tumors or Tumours or Cancer or Cancers oncolog*).ab,hw,ti.	82.907
8	6 or 7	84.215
9	exp Test Construction/ or exp Psychometrics/ or exp Rating Scales/ or exp Measurement/ or exp Test Validity/ or exp Psychological Assessment/	479.999
10	exp Questionnaires/ or exp Surveys/ or exp "Quality of Life"/ or exp Test Sensitivity/ or exp Self-Report/	98.977
11	(test* or scale* or questionnaire* or instrument* or tool* or inventor* or Thermometer* or intervent* or surve* or screen*).ab,hw,ti.	2.033.641
12	9 or 10 or 11	2.151.250
13	(develop* or validat* or construct* or "psychometric properties" or "diagnostic accuracy" or screen* or adapt* or translat* or evaluat* or assess* or detect*).ab,hw,ti.	2.578.264
14	spain.mp. or (spanish or espa?a or espa?ol).ab,hw,ti.	51.882
15	5 and 8 and 12 and 13 and 14	41
16	limit 15 to ("300 adulthood <age 18 yrs and older>" or 320 young adulthood <age 18 to 29 yrs> or 340 thirties <age 30 to 39 yrs> or 360 middle age <age 40 to 64 yrs> or "380 aged <age 65 yrs and older>" or "390 very old <age 85 yrs and older>")	33

## CINAHL (EBSCO)

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
S1	TI ( emotion* or psycholog* OR psychosocial ) AND TI ( distress* or stress* )	8264
S2	AB ( emotion* or psycholog* OR psychosocial ) AND AB ( distress* or stress* )	55.773
S3	MH "Psychological Distress"	2219
S4	S1 OR S2 OR S3	59.167
S5	TI ( Neoplasia or Neoplasias or Neoplasm* or Tumor or Tumour or Tumors or Tumours or Cancer or Cancers oncolog* ) OR AB ( Neoplasia or Neoplasias or Neoplasm* or Tumor or Tumour or Tumors or Tumours or Cancer or Cancers oncolog* ) OR MH Neoplasms	581.454
S6	MH ( "manifest anxiety scale" or "test anxiety scale" ) OR MH ( "surveys and questionnaires" or "health care surveys" or "health surveys" or "patient health questionnaire" or "self report" ) OR MH ( "Sensitivity and Specificity" OR "Psychometrics" or "Interviews as Topic" or "Patient Outcome Assessment" or "Interview, Psychological" ) OR TI ( test* or scale* or questionnaire* or instrument* or tool* or inventor* or Thermometer* or intervent* or surve* ) OR AB ( test* or scale* or questionnaire* or instrument* or tool* or inventor* or Thermometer* or intervent* or surve* )	1.792.500
S7	TI ( develop* or validat* or construct* or "psychometric properties" or "diagnostic accuracy" or screen* or adapt* or translat* or evaluat* or assess* or detect* ) OR AB ( develop* or validat* or construct* or "psychometric properties" or "diagnostic accuracy" or screen* or adapt* or translat* or evaluat* or assess* or detect* )	2.398.806
S8	TI ( spain or spanish or espa?a or espa?ol ) OR AB ( spain or spanish or espa?a or espa?ol ) OR MH spain	41.627
S9	S4 AND S5 AND S6 AND S7 AND S8 Limitadores - Grupos de edad: Adult: 19-44 years, Middle Aged: 45-64 years, Aged: 65+ years, Aged, 80 and over	50

## Web of Science

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
#1	TS=(emotion* or psychologic* or psychosocial) NEAR/5 (distress* or stress*) OR TI=((emotion* or psychologic* or psychosocial) NEAR/5 (distress* or stress*))	93.544
#2	(TS=(Neoplasia or Neoplasias or Neoplasm* or Tumor or Tumour or Tumors or Tumours or Cancer or Cancers oncolog*)) OR TI=(Neoplasia or Neoplasias or Neoplasm* or Tumor or Tumour or Tumors or Tumours or Cancer or Cancers oncolog*)	3.903.248

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
#3	(TS=("quality of life" OR Psychometr* OR Self?Report OR Measure* OR test* or scale* or questionnaire* or instrument* or tool* or inventor* or Thermometer* or intervent* or surve* or screen*)) OR TI=("quality of life" OR Psychometr* OR Self?Report OR Measure* OR test* or scale* or questionnaire* or instrument* or tool* or inventor* or Thermometer* or intervent* or surve* or screen*)	18.101.581
#4	(TS=(develop* or validat* or construct* or "psychometric properties" or "diagnostic accuracy" or screen* or adapt* or translat* or evaluat* or assess* or detect*)) OR TS=(develop* or validat* or construct* or "psychometric properties" or "diagnostic accuracy" or screen* or adapt* or translat* or evaluat* or assess* or detect*)	22.935.959
#5	(TS=(spain or spanish or espa?a or espa?ol)) OR TI=(spain or spanish or espa?a or espa?ol)	380.253
#6	#1 AND #2 AND #3 AND #4 AND #5	97

## BÚSQUEDA DE LITERATURA GRIS

Dialnet tesis doctorales

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
#1	Malestar emocional cáncer	29
#2	distrés emocional cáncer	17
#3	distrés psicológico cáncer	12
	Eliminados duplicados	29

Teseo

	<b>Términos de búsqueda</b>	<b>N.º de ítems</b>
#1	Malestar emocional cáncer	2
#2	distrés emocional cáncer	0
#3	distrés psicológico cáncer	0

El resultado de todas estas búsquedas se volcó en el gestor de referencias bibliográficas EndNote 20.2, con el fin de eliminar los duplicados de cada una de ellas.

## Anexo B. Tablas de evidencia

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
Galdón, 2008 (90)	<p><b>Objetivos:</b> Estudiar las propiedades psicométricas y la estructura factorial del Inventario Breve de Síntomas-18 (Brief Symptom Inventory-18, BSI-18) en una muestra de mujeres con cáncer de mama en España.</p> <p><b>Localización:</b> Fundación Instituto Valenciano de Oncología (FIVO, Valencia).</p> <p><b>Periodo de realización:</b> No se explicita.</p>	<p><b>Población:</b> La población son mujeres adultas con cáncer de mama. Las participantes elegibles deben ser mujeres que cumplan los siguientes requisitos: (a) tener 18 años o más; (b) haber recibido un diagnóstico primario de cáncer de mama; (c) no estar participando en ningún ensayo clínico; (d) proporcionar el consentimiento informado.</p> <p><b>Prueba a estudio:</b> La herramienta a analizar es el Inventario Breve de Síntomas18 (BSI18), concretamente la versión española publicada por Derogatis (2000).</p> <p><b>Prueba de comparación:</b> No se evalúa la precisión diagnóstica de la herramienta, por lo que no se emplea ninguna prueba de comparación.</p> <p><b>Resultados analizados:</b> Las variables de resultado analizadas son los índices de bondad de ajuste de la estructura factorial: RMSEA; Non-Normed Fit Index (NNFI); Robust Comparative Fit Index (RCFI); Chi cuadrado de Satorra Bentler (<math>S\chi^2</math>; <math>S-B\chi^2/gl</math>); Akaike Information Criterion (AIC); y la consistencia interna (alfa de Cronbach).</p>	<p><b>Número de participantes:</b> 175 mujeres que habían sido recientemente diagnosticadas de cáncer de mama y acuden a la FIVO (Valencia). Edad media = 52,7 años (dt = 10,12) Rango de edad = de 27 a 70 años. El 78 % está casada o vive con una pareja estable. Con respecto a su situación laboral, el 60 % de las participantes son amas de casa, mientras que el 36 % trabaja fuera del ámbito doméstico. Con relación al nivel educativo, el 84 % ha completado únicamente estudios primarios. Por último, la mayoría están en estadios tempranos de cáncer de mama (92 % en estadios I y II), y solo el 8 %, en estadios avanzados (estadio III).</p> <p><b>Criterios prueba diagnóstica positiva:</b> Se considera que una persona presenta malestar emocional si su puntuación total (calculada como puntuación estandarizada T) es igual o superior a 63 en el Índice Global de Severidad, o en al menos dos de las dimensiones.</p> <p><b>Criterios prueba de comparación positiva:</b> No se evalúa la precisión diagnóstica de la herramienta, por lo que no se emplea ninguna prueba de comparación.</p>	<p><b>Estructura factorial:</b> El Análisis Factorial Confirmatorio (AFC) con el método de máxima verosimilitud informa de una solución jerárquica con tres factores de primer orden (somatización, depresión y ansiedad) y un factor de segundo orden (GSI o puntuación total). Los índices de bondad de ajuste son adecuados para dicho modelo: <math>S-B\chi^2 = 148,67</math>; <math>S-B\chi^2/gl: 1,16</math>; Robust NNFI: 0,93; RCFI: 0,94; RMSEA (IC 90 %) = 0,03 (0,00 - 0,05); AIC = -107,33.</p> <p><b>Consistencia interna:</b> La consistencia interna tanto de la puntuación total (alfa de Cronbach (<math>\alpha</math>) = 0,82) como de las subescalas (<math>\alpha</math> somatización = 0,62; <math>\alpha</math> depresión = 0,73; <math>\alpha</math> ansiedad = 0,72) fue adecuada.</p>	<p>Las autoras del estudio concluyen que el BSI-18 es una herramienta válida y fiable.</p> <p>Cuenta con una estructura trifactorial (somatización, depresión y ansiedad) e informa de índices de bondad de ajuste adecuados para esa solución factorial.</p> <p>También se confirma que la herramienta presenta una consistencia interna adecuada.</p> <p>Las autoras afirman que, en el momento de publicación del artículo, no hay ningún otro instrumento de cribado de malestar emocional en población con cáncer que cuente con una mejor validación, por lo que creen que puede considerarse un potencial instrumento para elegir tanto en el ámbito clínico como en el de la investigación.</p>	<p>Riesgo de sesgo incierto con respecto a la selección de pacientes. No genera preocupaciones en cuanto a la aplicabilidad.</p>

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad
Martínez, 2019 (92)	<p><b>Objetivos:</b> Examinar la estructura factorial del Inventario Breve de Síntomas 18 (Brief Symptom Inventory-18, BSI18) en una muestra amplia y heterogénea de población oncológica adulta; proporcionar información sobre la validez convergente y la consistencia interna del instrumento; y determinar la invarianza factorial del BSI-18 para hombres y mujeres.</p> <p><b>Localización:</b> Fundación Instituto Valenciano de Oncología (Valencia, España).</p> <p><b>Periodo de realización:</b> Entre mayo de 2015 y julio de 2017.</p>	<p><b>Población:</b> La población son pacientes oncológicos en edad adulta. Los y las participantes elegibles son pacientes que cumplan los siguientes requisitos: (a) tener 18 años o más; (b) tener un diagnóstico de cáncer; (c) poder proporcionar el consentimiento informado.</p> <p><b>Prueba a estudio:</b> La herramienta a analizar es el Inventario Breve de Síntomas-18 (BSI-18), concretamente la versión española publicada por Derogatis (2013). Se trata de un cuestionario autoadministrado que consta de 18 ítems, con una escala de respuesta tipo Likert de 5 puntos que refleja el malestar durante la semana anterior. El instrumento aporta una puntuación total, y por subescalas (somatización, depresión y ansiedad).</p> <p><b>Otras pruebas:</b> Para evaluar la validez convergente del BSI18, se utilizó la Lista de Problemas (NCCN®, 2017), que consiste en un listado de 36 problemas experimentados de forma habitual por pacientes oncológicos, agrupados en 5 categorías: prácticos, familiares, emocionales, físicos y espirituales o religiosos. Han de responder si han experimentado o no alguno de los problemas durante la semana anterior.</p> <p><b>Resultados analizados:</b> Con respecto a la estructura factorial, las variables de resultado analizadas son los índices de bondad de ajuste de los modelos factoriales: Chi cuadrado de Satorra Bentler (S-B<math>\chi^2</math>); Comparative Fit Index (CFI); Root Mean Square Error of Approximation (RMSEA); Non-Normed Fit Index (NNFI), y Standardized Root Mean Square Error of the Residuals (SRMR). La fiabilidad del BSI-18 se midió mediante el alfa de Cronbach (consistencia interna) junto con la fiabilidad compuesta corregida por el error de las covarianzas (CR). La validez convergente del BSI-18 se evaluó a través de coeficientes de correlación entre el BSI-18 y la Lista de Problemas de la NCCN®.</p>	<p><b>Número de participantes:</b> La muestra final está formada por 1183 pacientes oncológicos (65 % mujeres). Edad media = 57,71 años (dt = 13,09; rango de edad = 19-94 años). La mayoría de las y los participantes convive con su pareja (73 %) y ha completado los estudios primarios (87 %). El tipo de cáncer primario más frecuente fue el de mama (40 %), seguido por el ginecológico (15 %) y el de próstata (12 %), otros tipos de cánceres primarios localizados, incluyendo respiratorio (7 %), digestivo (7 %), melanoma (6 %), y otros tipos de cáncer (9 %). Alrededor del 50 % de los y las participantes estaba en un estadio no avanzado de la enfermedad (estadios I y II), el 20 % estaba en el estadio III y el 22 %, en el estadio IV. Por último, el 64 % de las personas participantes estaba recibiendo tratamiento médico en el momento del estudio: cirugía (10 %), tratamiento adyuvante (47 %) y hormonoterapia (7 %).</p> <p><b>Criterios prueba diagnóstica positiva:</b> Se considera que una persona presenta malestar emocional clínicamente significativo si su puntuación total (calculada como puntuación estandarizada T) es igual o superior a 63 en el Índice Global de Severidad, o en al menos dos de las dimensiones.</p> <p><b>Criterios prueba de comparación positiva:</b> No se evalúa la precisión diagnóstica de la herramienta, por lo que no se emplea ninguna prueba de comparación.</p>	<p><b>Estructura factorial:</b> Teniendo en cuenta las bases teóricas y el Análisis Factorial Confirmatorio (AFC), se seleccionó un modelo jerárquico con tres factores de primer orden (somatización, depresión y ansiedad) y un factor de segundo orden (puntuación total o GSI). Los índices de bondad de ajuste son: S-B<math>\chi^2</math> = 366,9; S-B<math>\chi^2</math> /gl = 2,82; CFI = 0,927; NNFI = 0,915; RMSEA (IC 90 %) = 0,039 (0,035 - 0,044); SRMR = 0,037.</p> <p><b>Consistencia interna:</b> La consistencia interna de la puntuación total fue adecuada (alfa de Cronbach (<math>\alpha</math>) = 0,91), así como la de cada una de las subescalas (<math>\alpha</math> somatización = 0,72; <math>\alpha</math> depresión = 0,83; <math>\alpha</math> ansiedad = 0,84). La fiabilidad compuesta corregida por el error de las covarianzas fue satisfactoria tanto para la puntuación total (CR total = 0,90) como para cada una de las subescalas (CR somatización = 0,69; CR depresión = 0,79; CR ansiedad = 0,79).</p> <p><b>Validez convergente:</b> La validez convergente del BSI18 se evaluó mediante los coeficientes de correlación entre el BSI-18 (Índice Global de Severidad -GSI- y subescalas) y las 5 categorías de la Lista de Problemas de la NCCN®. El GSI, la subescala de depresión y de ansiedad correlacionan en mayor medida con los problemas emocionales, seguido de los problemas físicos, mientras que la subescala de somatización muestra una correlación mayor con los problemas físicos y, en segundo lugar, con los emocionales.</p>	<p>Las autoras del estudio concluyen que el BSI-18 es un instrumento de referencia para evaluar el malestar emocional de las y los pacientes con cáncer. Cuenta con una estructura jerárquica con tres factores de primer orden (somatización, depresión y ansiedad) y un factor de segundo orden (GSI o puntuación total); e informa de índices de bondad de ajuste adecuados para esa solución factorial.</p> <p>Además, se confirma que la herramienta presenta una fiabilidad adecuada, tanto en términos de consistencia interna como de fiabilidad compuesta.</p> <p>Se confirma la validez convergente del Índice Global de Severidad del BSI-18, así como de las puntuaciones de las subescalas, a través de las correlaciones con las categorías de problemas emocionales y de problemas físicos de la Lista de Problemas de la NCCN®.</p>	<p>No genera preocupaciones respecto al riesgo de sesgo ni a la aplicabilidad del estudio.</p>

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad
<b>Calderon, 2020 (93)</b>	<p><b>Objetivos:</b> Evaluar la estructura factorial y las propiedades psicométricas del Inventario Breve de Síntomas-18 (Brief Symptom Inventory-18, BSI-18) en pacientes con cáncer; evaluar la invarianza de las puntuaciones del BSI-18 con respecto al sexo, la edad y la localización del tumor; valorar la validez de constructo de las puntuaciones; y evaluar su sensibilidad al cambio inducido por el tratamiento, entre personas con cáncer no avanzado y extirpado.</p> <p><b>Localización:</b> Estudio multicéntrico del grupo de Cuidados Continuos de la Sociedad Española de Oncología Médica (SEOM), llevado a cabo en 15 departamentos de oncología médica, en 10 comunidades autónomas de España.</p>	<p><b>Población:</b> La población consiste en pacientes con cáncer confirmados histológicamente, no avanzado y tratado quirúrgicamente, para quienes las directrices clínicas internacionales consideraban que el tratamiento adyuvante era una opción. Se excluyeron las y los pacientes que: a) eran menores de 18 años; b) habían recibido tratamiento con radio o quimioterapia preoperatoria, o solo con terapia hormonal o con radioterapia adyuvante sin quimioterapia; o c) tenían alguna enfermedad mental grave que les impidiera comprender el estudio.</p> <p><b>Prueba a estudio:</b> La herramienta a analizar es el Inventario Breve de Síntomas-18 (BSI-18), concretamente la versión publicada por Derogatis (2001). Este instrumento incluye 18 ítems divididos en tres dimensiones (somatización, depresión y ansiedad) y la puntuación total (el Índice Global de Severidad -GSI-) para evaluar el malestar psicológico general en poblaciones clínicas y comunitarias. Las y los participantes han de responder sobre cómo se han sentido en los últimos 7 días y cada ítem se califica en una escala de respuesta tipo Likert de 5 puntos, desde 0 (Nada) hasta 4 (Extremadamente).</p>	<p><b>Número de participantes:</b> La muestra está formada por 877 pacientes, con una media de edad de 36 años (dt = 13,4; rango entre 25 y 84 años); siendo el 59,7 % mujeres. En cuanto al lugar de residencia, el 40,5 % viven en una ciudad grande (&gt;100 000 habitantes); el 26,6 %, en una ciudad mediana (entre 10 000 y 100 000), y el 32,8 %, en una ciudad pequeña (&lt;10 000). En cuanto a las características clínicas, los tipos de cáncer más frecuentes fueron el de colon (47,4 %) y el de mama (31,4 %), en estadio III (55,4 %).</p> <p><b>Criterios prueba diagnóstica positiva:</b> Las puntuaciones brutas del BSI-18 se convierten en puntuaciones T estandarizadas que se caracterizan por una distribución con una media de 50 y una desviación estándar de 10. De acuerdo con el punto de corte recomendado por Derogatis (2001), se considera que una puntuación T ≥ 63 corresponde a una persona que presenta malestar emocional.</p> <p><b>Criterios prueba de comparación positiva:</b> No se evalúa la precisión diagnóstica de la herramienta, por lo que no se emplea ninguna prueba de comparación.</p>	<p><b>Estructura factorial:</b> Los resultados del AFE entre muestras fueron casi idénticos en todos los casos (los coeficientes de congruencia fueron siempre superiores a 0,96). Este análisis sugiere que la mejor solución factorial es una solución de segundo orden, con tres factores primarios fuertemente relacionados y un factor general de segundo orden (SRMRS = 0,034; RMSEA = 0,032; CFI = 0,997). En función de los resultados del AFE, se ajustó un modelo CFA de segundo orden con tres factores primarios, usando de nuevo un esquema de validación cruzada basado en muestras aleatorias independientes. Los resultados del modelo propuesto fueron prácticamente idénticos (<math>\chi^2/gf = 2,84</math>; RMSEA (IC 90 %) = 0,046 (0,041 - 0,051); CFI = 0,978).</p> <p><b>Consistencia interna:</b> La consistencia interna de la puntuación total fue adecuada (alfa de Cronbach (<math>\alpha</math>) = 0,91), así como la de cada una de las subescalas (<math>\alpha</math> somatización = 0,72; <math>\alpha</math> depresión = 0,83; <math>\alpha</math> ansiedad = 0,89).</p>	<p>Las autoras y los autores del estudio concluyen que el BSI-18 cuenta con una solución factorial que encaja con la estructura originalmente propuesta, de tres factores primarios y un factor general de segundo orden.</p> <p>También se confirma que las puntuaciones de la herramienta presentan una adecuada fiabilidad.</p> <p>Con respecto a la validez de constructo, el factor de depresión fue el que correlacionó más fuertemente con el apoyo social, mientras que las escalas de somatización y ansiedad también se asociaron negativamente con el apoyo social y todas las subescalas, aunque con tamaños de efecto menores.</p>	No genera preocupaciones respecto al riesgo de sesgo ni a la aplicabilidad del estudio.

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad
<p><b>Calderon, 2020 (93)</b></p>	<p><b>Periodo de realización:</b> De julio de 2015 a febrero de 2019</p>	<p><b>Otras pruebas:</b> Para evaluar la validez de constructo, se empleó la Escala Multidimensional de Apoyo Social Percibido (MSPSS). Se trata de una escala de autoinforme de 12 ítems que mide la percepción y la adecuación del apoyo social de tres fuentes: la familia, las amistades y las personas importantes. La puntuación total refleja el grado total de apoyo social que reciben los individuos. Los ítems se puntúan en una escala de respuesta tipo Likert de 7 puntos y la puntuación total va de 7 a 84. Las puntuaciones más altas indican un mayor apoyo percibido.</p> <p><b>Resultados analizados:</b> Con respecto a la estructura factorial, se calculan los índices de bondad de ajuste (SRMRS, RMSEA, CFI), tanto para la solución ofrecida por el AFE como por el AFC, basados en validaciones cruzadas en submuestras aleatorias e independientes.</p> <p>La consistencia interna del BSI-18 se midió mediante el alfa de Cronbach.</p> <p>La validez convergente del BSI-18 se evaluó a través de los coeficientes de correlación brutos, así como de los estandarizados, entre el BSI-18 (Índice Global de Severidad, GSI, y las 3 subescalas) y el MSPSS (las 3 subcategorías y el total).</p>		<p><b>Validez convergente:</b> La validez convergente se evaluó mediante los coeficientes de correlación brutos, así como los estandarizados, entre el BSI-18 (Índice de Severidad Global –GSI- y las 3 subescalas) y el MSPSS (las 3 subcategorías y el total). El patrón de correlaciones brutas es similar al de correlaciones estandarizadas, como cabría esperar. El factor de depresión (BSI-18) es el que correlacionó más fuerte con las escalas del MSPSS (correlación estandarizada depresión-familia= -0,31; depresión-amistades: -0,26; depresión-otras personas importantes: 0,28; depresión-total MSPSS: 0,36). Las escalas de somatización y ansiedad y el GSI total también correlacionan negativa y significativamente con todas las escalas del MSPSS y el MSPSS total, aunque los tamaños del efecto fueron menores.</p>		

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad
<b>Carvajal, 2013 (100)</b>	<p><b>Objetivos:</b> Replicar el estudio de Watanabe <i>et al.</i> (2011) utilizando la versión española del ESAS-r, con una muestra internacional de pacientes de lengua española. Además, para obtener la primera evidencia de validez de la nueva versión ESAS-r, sus propiedades psicométricas fueron evaluadas y comparadas con los datos de la versión original.</p> <p><b>Localización:</b> Departamentos de Oncología de hospitales terciarios (Hospital de la Universidad de Navarra, España, e Instituto Nacional del Cáncer, Guatemala).</p>	<p><b>Población:</b> La población objeto de estudio son las personas pacientes con cáncer avanzado. Todas eran mayores de 18 años, con una función cognitiva normal en el Mini-Mental State Examination (MMSE). Los y las pacientes con cáncer avanzado que estaban recibiendo radioterapia paliativa o quimioterapia también eran elegibles.</p> <p><b>Prueba a estudio:</b> La prueba diagnóstica de interés es la versión española del Edmonton Symptom Assessment System-Revised (ESAS-r). La versión revisada incluye clarificación del número de términos, cambio en el formato, la definición de un marco temporal específico, y un reordenamiento de los ítems. El ESAS-r consiste en la evaluación de la presencia de diez síntomas en el momento actual, mediante una escala numérica visual que va desde 0 a 10 (más un ítem abierto donde el o la paciente puede incluir algún otro problema).</p> <p><b>Prueba de comparación:</b> No se explora su precisión diagnóstica, por lo que no hay una prueba de comparación en ese sentido. No obstante, se comparan los resultados del ESAS-r con los de la antigua versión de la escala, el ESAS.</p>	<p><b>Número de participantes:</b> Un total de 14 pacientes con cáncer avanzado participaron en el estudio piloto del ESAS-r.  La muestra total de pacientes para el estudio psicométrico del ESAS-r fue de 66 pacientes (al menos 30 de cada centro participante). El 70 % eran mujeres, y la edad media fue de 54 años (en un rango de 18 a 84 años). El diagnóstico de cáncer primario más frecuente era el gastrointestinal (36 %), seguido del genitourinario (24 %). El tratamiento actual de cáncer más frecuente era la quimioterapia para el 51 % y los cuidados paliativos para el 39 %. El 47 % tenía estudios primarios, el 30 % estudios secundarios y el 23 % educación superior.</p> <p><b>Criterios prueba diagnóstica positiva:</b> No se describe qué se considera resultado negativo y positivo en el ESAS-r.</p> <p><b>Criterios prueba de comparación positiva:</b> No existe prueba de comparación, ya que no se busca informar de la precisión diagnóstica de la herramienta.</p>	<p><b>Consistencia interna:</b> La consistencia interna obtenida fue un coeficiente alfa de Cronbach de 0,86.  La correlación de Spearman entre el ítem de bienestar y el total de ítems individuales es de 0,67.</p> <p><b>Validez discriminante:</b> Se encontraron correlaciones negativas y significativas (desde <math>r_s = -0,38</math> a <math>-0,45</math>) entre las puntuaciones de los y las participantes con respecto al estatus funcional medido por el PPS y cinco ítems del ESAS-r (dolor, cansancio, somnolencia, pérdida de apetito y bienestar). No se encontraron correlaciones significativas entre el PPS y otros cinco ítems (náuseas, ansiedad, depresión, dificultad para dormir y dificultad para respirar).</p>	<p>Las autoras y los autores concluyen que el ESAS-r es un instrumento válido que mejora las propiedades psicométricas de las versiones previas de la herramienta, por lo que puede reemplazar a las versiones previas del instrumento. Se ha desarrollado la versión española del ESAS revisada, y ha sido percibida por parte de los y las pacientes como más fácil de entender y completar que el ESAS. Tanto los pacientes que alguna vez habían completado el ESAS como los que no prefieren el ESAS-r.</p> <p>La consistencia interna del ESAS-r fue mejor que la del ESAS, en la versión española. La correlación del ítem de sensación de bienestar con el global de los ítems fue moderada, sugiriendo que este ítem puede estar relacionado con la carga global de los síntomas.</p>	<p>Riesgo de sesgo incierto con respecto a la selección y el flujo de pacientes.  Genera preocupaciones sobre la aplicabilidad.</p>

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad
Carvajal, 2013 (100)	<p><b>Periodo de realización:</b></p> <p>De diciembre de 2010 a febrero de 2011.</p>	<p><b>Resultados analizados:</b></p> <p>Con respecto al estudio psicométrico de la escala, se estudia su consistencia interna (alfa de Cronbach). También se calculó la correlación de Speaaman del ítem de bienestar con el total de ítems individuales.</p> <p>La validez discriminante se calculó evaluando las correlaciones de la puntuación ESAS-r entre pacientes con diferente estado funcional (usando la PPS, Palliative Performance Scale) mediante la rho de Spearman.</p>			<p>Las puntuaciones en cinco ítems del ESAS-r correlacionaron negativa y significativamente con el estado funcional del paciente, medido con el PPS. Esto parece consistente con el hecho de que aquellos pacientes con puntuaciones más altas en síntomas del ESAS-r tengan peor estado funcional.</p>	

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad
<b>Martínez, 2006 (101)</b>	<p><b>Objetivos:</b> Determinar la capacidad de las Escalas Visuales Analógicas (EVA) de ansiedad y depresión como instrumentos de <i>screening</i> de estrés emocional.</p> <p><b>Localización:</b> Consultas externas del Servicio de Cuidados Paliativos del Instituto Catalán de Oncología (ICO).</p> <p><b>Periodo de realización:</b> De septiembre a noviembre de 2001.</p>	<p><b>Población:</b> La población objeto de estudio son pacientes con diagnóstico de cáncer avanzado o terminal a quienes atienden los equipos de cuidados paliativos. Como criterios de inclusión se determinaron: (1) pacientes con diagnóstico de cáncer; (2) edad mayor de 18 años; (3) pacientes que acuden por primera vez en la consulta externa del Servicio de Cuidados Paliativos del Instituto Catalán de Oncología (ICO). Como criterio de exclusión se determinó la presencia de un deterioro del estado general del paciente que, a criterio de los y las profesionales, no permitiera la evaluación y la valoración completa de los diferentes parámetros que se miden en el estudio.</p> <p><b>Prueba a estudio:</b> Las pruebas índice son tanto la Escala Visual Analógica (EVA) de ansiedad como la de depresión. Se entiende que las puntuaciones van de 0 a 10, pero no se especifica. También se calcula el Distrés Emocional Global (DEG), resultado de sumar las puntuaciones de las EVA de ansiedad y depresión.</p> <p><b>Prueba de comparación:</b> La Escala Hospitalaria de Ansiedad y Depresión (HADS), compuesta por 14 ítems y dos subescalas (ansiedad y depresión). No especifican la escala de medida de los ítems.</p> <p><b>Resultados analizados:</b> Se calcula la correlación (r de Pearson) entre las EVA de ansiedad y depresión y las correspondientes subescalas de ansiedad y depresión de la HADS.</p> <p>Se calcula el área bajo la curva para determinar el punto de corte óptimo de las EVA, y los valores de sensibilidad y especificidad.</p>	<p><b>Número de participantes:</b> La muestra final está compuesta por 53 pacientes. El 66 % eran hombres. Edad media = 63 años (dt = 12,59). Las neoplasias más frecuentes fueron digestivas (24,5 %), cabeza y cuello (17 %), y mama (15 %). Un 30,2 % presentaba historia previa de adicción al alcohol y un 24,5 % de los y las pacientes refería antecedentes de psicopatología.</p> <p><b>Criterios prueba diagnóstica positiva:</b> Uno de los objetivos del estudio es determinar el punto de corte de la escala adecuado, por lo que no se establece <i>a priori</i>.</p> <p><b>Criterios prueba de comparación positiva:</b> El punto de corte de la HADS se establece en 8 puntos en las subescalas y 16 en el total.</p>	<p><b>Validez convergente:</b> Las correlaciones que se calcularon fueron estadísticamente significativas (<math>p &lt; 0,001</math>). En concreto, la correlación entre la EVA ansiedad y la HADS ansiedad fue de 0,663; la correlación entre la EVA depresión y la HADS depresión fue de 0,576; y la correlación entre el DEG y la HADS total fue de 0,770.</p> <p><b>Precisión diagnóstica:</b> El área bajo la curva ROC se calculó para la puntuación del DEG, siendo de 0,870. El punto de corte 10 se consideró una combinación óptima de sensibilidad (80 %) y especificidad (93,9 %).</p>	<p>El equipo autor concluye que sus resultados apoyan el uso sistemático de las EVA de ansiedad y depresión en las consultas externas como un instrumento de <i>screening</i> para detectar el distrés emocional significativo y tributario de valoración específico por psicólogas y psicólogos, cuando la suma de las puntuaciones de EVA de ansiedad y depresión sea igual o superior a 10.</p> <p>Complementariamente, observan que las correlaciones mejoran entre las sumas de puntuaciones, lo que indica que al final de la vida puede resultar difícil establecer límites precisos entre los sentimientos que expresa el paciente, de forma que tiene más sentido hablar de DEG.</p>	No genera preocupaciones respecto al riesgo de sesgo ni a la aplicabilidad.

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad
López-Roig, 2000 (94)	<p><b>Objetivos:</b></p> <p>Realizar un estudio psicométrico de la Escala Hospitalaria de Ansiedad y Depresión (HADS) en población española de pacientes con cáncer. Concretamente, se plantea: comprobar la fiabilidad del instrumento para la población española de pacientes oncológicos; estudiar qué constructos miden las escalas, si son capaces de discriminar entre respuestas de ansiedad y depresión; establecer el grado de especificidad y de sensibilidad para los estados de ansiedad y depresión, determinando los puntos de corte para clasificar niveles de respuesta; establecer las variaciones paramétricas en función del género y comprobar si otras variables sociodemográficas y clínicas introducen cambios a tener en cuenta en la utilización de las escalas, así como su utilidad para medir respuestas emocionales situacionales puntuales, durante períodos de la enfermedad.</p>	<p><b>Población:</b></p> <p>La población objeto de estudio es la población española de pacientes con cáncer.</p> <p><b>Prueba a estudio:</b></p> <p>La Escala Hospitalaria de Ansiedad y Depresión (HADS, Zigmond y Snaith, 1983) utilizada es la versión hispana remitida por los autores originales. Los autores y las autoras de este estudio modificaron su diseño para permitir su administración tanto en formato autoaplicado como de entrevista. Este instrumento es una medida situacional de ansiedad y depresión. No se trata de una medida puntual, sino que se aplica con relación a un período actual (la última semana). En el caso de pacientes oncológicos en fase de seguimiento, la respuesta emocional en la última semana puede estar contaminada por la anticipación de la visita para la revisión, por lo que en estos casos modifican el período temporal de referencia pidiendo al paciente que conteste: ¿Cómo se siente últimamente?, al margen de los días previos en que haya estado pensando en la visita. Consta de 14 ítems. Los ítems impares configuran la subescala de ansiedad y su escala de respuesta se puntúa de 3 a 0. Los pares conforman la subescala de depresión y se puntúan de 0 a 3. La puntuación total en cada subescala se obtiene sumando las puntuaciones de los ítems correspondientes, con un rango cada una de 0 a 21. En ambos casos, a mayor puntuación, mayor nivel de ansiedad o depresión.</p> <p><b>Prueba de comparación:</b></p> <p>Se utilizan dos instrumentos estándar. En el caso de la ansiedad, el Inventario de Ansiedad Estado-Rasgo (subescala de ansiedad-estado) (STAI-S; Spielberg, 1972; ed. Tea, 1982). Consta de 20 ítems con cuatro categorías de respuestas: casi nunca, a veces, a menudo y casi siempre.</p>	<p><b>Número de participantes:</b></p> <p>La muestra estudiada incluye 484 pacientes, 336 mujeres (69 %) y 148 hombres (31 %). Los y las pacientes se entrevistaron cuando acudían a las consultas externas: 139 pacientes (28,7 %); al hospital de día: 131 (27,1 %); o a la planta hospitalaria: 112 (23,1 %); mientras que 102 pacientes (21,1 %) se entrevistaron en la consulta de psicología. Edad media = 55,7 años (dt = 13,3); la mayoría están casados o casadas o viven en pareja (75 %); tienen estudios primarios (32,4 %) o leen y escriben (28,9 %). Del 80 % de las personas enfermas que informó acerca de sus ingresos mensuales, más de un tercio se sitúa en el nivel económico más bajo (33,3 %) y prácticamente el resto entre los niveles medio-bajo y medio-alto (43,8 %). Un 50 % de las mujeres son amas de casa y, de ellas, la mayoría mantiene su actividad sin cambios (56 %) o con alguna limitación (34 %). En cuanto a la localización del tumor primario, las mujeres con cáncer de mama superan el 50 % de la muestra total. El tiempo medio desde el diagnóstico es de 2,3 años (dt = 2,97). Los y las pacientes se encuentran en diferentes fases de la enfermedad, siendo mayoritario el porcentaje de pacientes en seguimiento (42,4 %), sobre todo por las mujeres con cáncer de mama, y en tratamiento (39 %). Un 37,1 % de los y las pacientes recibe tratamiento de quimioterapia, mientras que a un 36,1 % no se le aplica ningún tratamiento.</p> <p><b>Criterios prueba diagnóstica positiva:</b></p> <p>Este estudio explora los puntos de corte óptimos de la escala HADS, por lo que en un principio no están definidos.</p>	<p><b>Estructura factorial:</b></p> <p>El AFE recomienda una solución bifactorial (depresión y ansiedad), mostrando un índice KMO de 0,935 y un valor de <math>\chi^2 = 2955,6</math> (<math>p &lt; 0,001</math>) en el test de Bartlett, lo que indica su adecuación.</p> <p><b>Consistencia interna:</b></p> <p>Se calculó la consistencia interna de las dos subescalas (alfa de Cronbach – HADS-A = 0,85; alfa de Cronbach – HADS-D = 0,87). Se empleó el método test-retest, siendo la correlación de Spearman significativa (correlación test-retest HADS-A = 0,65; HADSD = 0,64).</p> <p><b>Validez:</b></p> <p>Las subescalas de ansiedad y depresión de la HADS correlacionan positiva y significativamente con las Escalas de ansiedad y depresión, respectivamente (<math>r</math> HADS-A – STAI-S = 0,71; <math>r</math> HADS-D – SDS = 0,76).</p> <p><b>Precisión diagnóstica:</b></p> <p>Se indica la sensibilidad y la especificidad de cada subescala de la HADS.</p> <p>– HADS-A (<math>n = 140</math>): punto de corte 9: (sensibilidad = 0,81; especificidad = 0,77);</p> <p>– HADS-D (<math>n = 87</math>): punto de corte 4: (sensibilidad = 0,80; especificidad = 0,81).</p>	<p>Las conclusiones de los autores y las autoras indican que la escala HADS puede utilizarse como medida de evaluación de ansiedad y depresión en población española de pacientes oncológicos de manera válida y fiable.</p> <p>Cuenta con una estructura bifactorial (ansiedad y depresión).</p> <p>También se confirma que la herramienta presenta una consistencia interna adecuada en cada una de sus subescalas.</p> <p>La fiabilidad test-retest presenta valores aceptables. Las diferencias significativas obtenidas entre ambas medidas pueden reflejar una situación emocional diferente de los y las pacientes entre la primera y la segunda medida. Estos resultados apoyan la estabilidad de la HADS, pero, a su vez, la capacidad de ser sensible a los cambios emocionales, a diferencia de lo que ocurriría con una medida de rasgo.</p>	<p>Riesgo de sesgo incierto con respecto a la selección de pacientes, la prueba índice y de referencia; riesgo de sesgo alto respecto al flujo y tiempos.</p> <p>Genera preocupaciones sobre la aplicabilidad de la prueba de referencia.</p>

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad
López-Roig, 2000 (94)	<p><b>Localización:</b> 7 hospitales o centros sitios en el Estado español.</p> <p><b>Periodo de realización:</b> No se explicita.</p>	<p>Para la depresión, se utiliza la Escala de depresión de Zung (SDS; Zung, 1965, adaptación española de Conde, Escribá e Izquierdo, 1970; Conde y De Esteban, 1976). Esta escala evalúa la intensidad de los síntomas cognitivos, afectivos psicomotores y fisiológicos asociados a la depresión, mediante 20 ítems basados en criterios diagnósticos comunes del trastorno depresivo. Las respuestas utilizan escalas de frecuencia.</p> <p><b>Resultados analizados:</b> Con respecto a la estructura factorial, se realizó un AFE, y se calcula el índice KMO y la prueba de esfericidad de Bartlett.</p> <p>Para examinar la consistencia interna, se utilizó la prueba alfa de Cronbach (correlación inter-elementos promedio), y se empleó el método test-retest para evaluar si la escala tiene estabilidad temporal, utilizando la correlación de Spearman y la prueba de los rangos con signo de Wilcoxon para el análisis de diferencias.</p> <p>La validez concurrente se exploró a través de las correlaciones entre las subescalas de ansiedad y depresión de la HADS con el STAI-S y el SDS, respectivamente. El coeficiente de correlación se calculó usando el método de Spearman.</p> <p>Estas mismas escalas (STAI-S y SDS) proporcionaron los valores de referencia para obtener los parámetros de sensibilidad y especificidad y los correspondientes puntos de corte.</p>	<p><b>Criterios prueba de comparación positiva:</b> En el caso de la escala de ansiedad STAI-S (Spielberg, 1972; ed. Tea, 1982), las puntuaciones de corte que determinan al sujeto como persona ansiosa en esta versión son 31 para los varones y 36 para las mujeres, correspondiendo a las medias de la población normal más una desviación típica (Percentil &gt;80).</p> <p>En el caso de la escala SDS (Zung, 1965, adaptación española de Conde, Escribá e Izquierdo, 1970; Conde y De Esteban, 1976), se consideran como casos aquellas personas cuyas puntuaciones directas superan o son iguales a 52, incluyendo a las y los pacientes que se clasifican según criterio clínico en el rango considerado como depresión media o grave (52-100).</p>		<p>Con respecto a los puntos de corte óptimos para considerar que la o el paciente oncológico está padeciendo algún grado de impacto emocional, se establece el punto de corte 9 para la subescala de ansiedad (10 en mujeres), y 4 para la subescala de depresión.</p> <p>El equipo autor apoya el uso de las dos subescalas por separado, como medidas diferentes de ansiedad y depresión, aunque puedan coexistir en el tiempo, y no como una medida global de distrés.</p>	

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad
<b>Costa, 2009 (95)</b>	<p><b>Objetivos:</b> Determinar la dimensionalidad de la HADS y su índice de fiabilidad, además de calcular los puntos de corte óptimos que permitan detectar probable sintomatología (ansiedad y depresión) clínicamente significativa en una muestra heterogénea de pacientes oncológicos ambulatorios, utilizando como criterio externo la entrevista psiquiátrica SCID-I, en la versión del DSM-IV (First <i>et al.</i>, 1999).</p> <p><b>Localización:</b> Hospital Duran y Reynals (Barcelona), perteneciente al Institut Català d'Oncologia.</p> <p><b>Periodo de realización:</b> De octubre de 2005 a junio de 2006.</p>	<p><b>Población:</b> La población son pacientes adultos oncológicos ambulatorios. Los criterios de inclusión fueron: a) ser paciente ambulatorio con diagnóstico de enfermedad oncológica; y b) tener una edad comprendida entre los 18 y los 80 años. Los criterios de exclusión fueron: a) presentar trastorno psicótico; b) tener una capacidad funcional menor del 40 % (según Karnofsky Performance Status); c) presentar un deterioro cognitivo que pudiera interferir en la comprensión de las pruebas administradas.</p> <p><b>Prueba a estudio:</b> La prueba diagnóstica de interés es la Escala Hospitalaria de Ansiedad y Depresión (Zigmond &amp; Snaith, 1983). Es un autoinforme con 14 ítems puntuados de 0 a 3. Con esta escala se evalúa la sintomatología de ansiedad y depresión con 7 ítems cada una. La versión española de la HADS fue adaptada y validada en pacientes con trastornos psiquiátricos (Tejero <i>et al.</i>, 1986) y ha sido validada en población oncológica, obteniendo buenos parámetros de fiabilidad (López Roig <i>et al.</i>, 2000).</p>	<p><b>Número de participantes:</b> La muestra final estuvo compuesta por 234 pacientes oncológicos ambulatorios, que acudían al hospital de día para recibir tratamiento de quimioterapia o asistían a las consultas en las unidades funcionales. Del total, 121 eran mujeres (51,7 %). Edad media = 55,24 años (dt = 11,5), dentro de un rango de 19 a 80 años. A nivel educativo, el 61,1 % cuenta con estudios primarios. Con respecto al estado civil, el 78,6 % está casada o casado o con pareja.</p> <p><b>Criterios prueba diagnóstica positiva:</b> Dado que uno de los objetivos del trabajo es calcular los puntos de corte óptimos de la HADS en pacientes oncológicos ambulatorios, no se define <i>a priori</i> el valor umbral o punto de corte.</p> <p><b>Criterios prueba de comparación positiva:</b> En la entrevista psiquiátrica, un resultado positivo se considera aquellos participantes que cuenten con sintomatología depresiva, de ansiedad o mixta (ansiedad-depresión). Se consideran negativos las y los pacientes sin trastorno psiquiátrico.</p>	<p><b>Estructura factorial:</b> El AFE evidencia la existencia de dos factores (depresión y ansiedad), y confirma la estructura bidimensional propuesta originariamente. El índice KMO es de 0,91 y la prueba de esfericidad de Bartlett obtiene un valor de <math>\chi^2 = 1242,86</math> (<math>p &lt; 0,001</math>).</p> <p><b>Consistencia interna</b> Se calculó la consistencia interna de las dos subescalas (alfa de Cronbach- HADS-A = 0,82; <math>\alpha</math> HADS-D = 0,84).</p> <p><b>Precisión diagnóstica:</b> – Para malestar emocional, utilizando la puntuación total en la HADS, el área bajo la curva ROC muestra una buena discriminación entre caso clínico y no clínico (área = 0,81; IC 95 % = 0,74 - 0,87). Se sugiere una puntuación de corte con valor 10, con una sensibilidad del 74 %, una especificidad del 72,7 %, un valor predictivo positivo del 55,1 % y un valor predictivo negativo del 86 %. – Para sintomatología depresiva, utilizando la subescala HADS-D, el área bajo la curva ROC muestra una buena discriminación entre caso clínico y no clínico (área = 0,74; IC 95 % = 0,65 - 0,83). Se sugiere una puntuación de corte con valor 4, con una sensibilidad del 73,5 %, una especificidad del 61,5 %, un valor predictivo positivo del 24 % y un valor predictivo negativo del 93,1 %.</p>	<p>El equipo autor del estudio concluye que la escala HADS cuenta con una estructura factorial bidimensional (ansiedad y depresión).</p> <p>Se hallan valores altos en consistencia interna tanto para la dimensión de ansiedad como de depresión.</p> <p>El punto de corte establecido como probable caso clínico fue 4 para la subescala de depresión, 8 para la subescala de ansiedad y 10 para la puntuación total.</p> <p>Los resultados del estudio enfatizan las características de la HADS como instrumento de cribado para evaluar cuantitativamente la presencia de probable sintomatología ansiosa o depresiva clínicamente significativa.</p>	<p>Riesgo de sesgo incierto sobre la selección de pacientes y la prueba de referencia. No genera preocupaciones en cuanto a la aplicabilidad.</p>

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad
Costa, 2009 (95)		<p><b>Prueba de comparación:</b> La prueba de comparación es la Entrevista clínica estructurada para los trastornos del eje I del DSM-IV (SCID-I; APA, 1994). La entrevista se realiza en una única sesión con una duración aproximada de 45 a 90 minutos. Los diagnósticos obtenidos con la entrevista psiquiátrica se agrupan, según la sintomatología, en cuatro categorías: a) sintomatología depresiva; b) sintomatología de ansiedad; c) sintomatología mixta (ansiedad-depresión); d) pacientes sin trastorno psiquiátrico.</p> <p><b>Resultados analizados:</b> Para comprobar la dimensionalidad de la escala, se realizó un AFE y se calculó la prueba de esfericidad de Bartlett y el índice Kaiser-Meyer-Olkin (KMO). Para medir la consistencia interna, se calculó el coeficiente alfa de Cronbach para cada dimensión del cuestionario. Se calculó una medida global de la exactitud de la prueba para el conjunto de todos los posibles puntos de corte mediante la curva ROC (Doménech, 1999). También se calcularon los índices de sensibilidad y especificidad, VPP, VPN para la escala en su conjunto y para cada subescala.</p>		<p>– Para sintomatología ansiosa, utilizando la subescala HADS-A, el área bajo la curva ROC muestra buena discriminación entre caso clínico y no clínico (área = 0,78; IC 95 %: 0,66 - 0,90). Se sugiere una puntuación de corte con valor 8, con una sensibilidad del 71,4 %, una especificidad del 76,5 %, un valor predictivo positivo del 23,1 % y un valor predictivo negativo del 96,4 %.</p>		

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
<b>Martínez 2012 (96)</b>	<p><b>Objetivos:</b></p> <p>A) Determinar si la puntuación total de la HADS puede utilizarse como una medida fiable de malestar emocional general en una muestra heterogénea amplia de pacientes oncológicos; b) comparar —usando el Análisis Factorial Confirmatorio— el modelo bifactorial de la HADS originalmente propuesto y un modelo jerárquico (dos escalas y su combinación); c) determinar —usando el análisis de la curva ROC— la sensibilidad y la especificidad de varios puntos de corte para detectar casos clínicos de malestar emocional general; y d) recomendar un punto de corte óptimo para el total de la escala para usar en población oncológica.</p>	<p><b>Población:</b></p> <p>La población objeto de estudio son pacientes oncológicos en edad adulta. Los criterios de inclusión fueron: a) tener 18 años o más; b) tener un diagnóstico de cáncer; y c) ser capaz de proporcionar el consentimiento informado.</p> <p><b>Prueba a estudio:</b></p> <p>La HADS (Escala Hospitalaria de Ansiedad y Depresión) es una escala de autoinforme especialmente diseñada para evaluar el malestar emocional en enfermedades físicas. Incluye 14 ítems, la mitad miden la ansiedad y la otra mitad, la depresión. Los ítems se puntúan en una escala de respuesta tipo Likert de 4 puntos. En este estudio, se ha utilizado la traducción española publicada por Caro e Ibáñez (1992). Esta versión ha mostrado índices adecuados de fiabilidad y validez en diferentes poblaciones (general, psiquiátrica y clínica).</p>	<p><b>Número de participantes:</b></p> <p>De los y las 945 pacientes que se abordaron, el 6 % declinó participar en el estudio. La muestra final estuvo formada por 892 pacientes oncológicos, cuyo rango de edad oscilaba entre 24 y 93 años, siendo la media de 57 años (dt = 13,3). El 70 % eran mujeres; la mayoría de las personas de la muestra estaban casadas o viviendo con una pareja estable (72 %), y tenían completados, al menos, los estudios primarios (85 %). En cuanto a la situación laboral, el 32 % de los y las participantes se había jubilado, el 30 % no tenía empleo o estaba en situación de baja laboral, el 19 % eran amas de casa y el 17 % trabajaban fuera de casa. Con respecto a las características médicas, el tipo de cáncer más frecuente era el de mama (42,60 %), en estadio II (28,80 %), y en fase de tratamiento (61,90 %).</p> <p><b>Criterios prueba diagnóstica positiva:</b></p> <p>Uno de los objetivos del estudio es determinar el punto de corte adecuado de la escala, por lo que no se establece a priori.</p>	<p><b>Estructura factorial:</b></p> <p>El AFC confirmó que el modelo jerárquico de tres factores (subescalas de ansiedad y depresión y escala total) es el que presenta unos mejores índices de bondad de ajuste (RMSEA (IC 90 %) = 0,04 (0,03 - 0,05); RCFI: 0,97; IFI: 0,97; Robust NNFI: 0,96; S-B<math>\chi^2</math>: 177,23; SB<math>\chi^2</math>/gl: 2,36; AIC: 27,23).</p> <p><b>Fiabilidad:</b></p> <p>El alfa de Cronbach de las subescalas de ansiedad y depresión fue de 0,83 en ambos casos, y 0,89 en la escala total (malestar emocional).</p> <p><b>Precisión diagnóstica:</b></p> <p>El análisis del área bajo la curva ROC indica que la puntuación total de la HADS tiene una excelente utilidad diagnóstica en relación con el BSI-18 (área = 0,95; IC 95 %: 0,94 - 0,97). El punto de corte 16 se consideró como la combinación óptima de sensibilidad (80 %) y especificidad (92 %). El índice de exactitud, corregido por la prevalencia, fue de 0,90.</p>	<p>Las autoras del estudio concluyen que, teniendo en cuenta el análisis factorial confirmatorio realizado y los índices de bondad de ajuste de los modelos, la solución jerárquica trifactorial es la estructura más apropiada para usar en población oncológica.</p> <p>El modelo jerárquico obtenido cuenta con un aval metodológico riguroso y puede resultar más útil que el modelo bifactorial original en programas de <i>screening</i> de malestar emocional, ya que permite una puntuación total.</p>	<p>No genera preocupaciones respecto al riesgo de sesgo ni a la aplicabilidad.</p>

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
Martínez 2012 (96)	<p><b>Localización:</b> Fundación Instituto Valenciano de Oncología (FIVO).</p> <p><b>Periodo de realización:</b> Desde octubre de 2010 a abril de 2012.</p>	<p><b>Prueba de comparación:</b> La prueba utilizada como comparación es el Inventario Breve de Síntomas-18 (BSI-18; Derogatis, 2000), una medida de autoinforme del malestar emocional. Los 18 ítems de los que está compuesta se califican en una escala de respuesta tipo Likert de 5 puntos. El instrumento permite calcular una puntuación total que consiste en la suma de los ítems, y una puntuación para cada subescala de 6 ítems (ansiedad, depresión y somatización). Para este estudio, se empleó la traducción española publicada por Derogatis (2000), y se consideró únicamente la puntuación total como medida de malestar emocional.</p> <p><b>Resultados analizados:</b> Se llevó a cabo un AFC con el método de máxima verosimilitud. Se calcularon los índices de bondad de ajuste de los modelos: RMSEA (IC 90 %); NNFI; RCFI; IFI. También se calculó el Chi cuadrado de Santorra-Bentler (<math>S-B\chi^2</math>; <math>S-B\chi^2/df</math>) y el Akaike Information Criterion (AIC).</p> <p>Se calculó la consistencia interna con el coeficiente alfa de Cronbach.</p> <p>Se realizó un análisis de la curva ROC para explorar el punto de corte óptimo de la HADS como medida de malestar emocional, calculando la sensibilidad, la especificidad y un índice de exactitud (corregido por la prevalencia).</p>	<p><b>Criterios prueba de comparación positiva:</b> Las puntuaciones brutas del BSI-18 se convierten en puntuaciones T estandarizadas. Las personas con una puntuación <math>T \geq 63</math> en la puntuación total, o en al menos dos de las tres subescalas, se consideran como casos.</p>		<p>Tanto la escala total como las dos subescalas muestran una consistencia interna adecuada.</p> <p>El punto de corte óptimo establecido fue 16. Sin embargo, también señalan que podría emplearse cualquier punto de corte entre 13 y 16, debido a los adecuados niveles de sensibilidad y especificidad asociados.</p>	

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
<b>Martínez, 2013 (97)</b>	<p><b>Objetivos:</b> Los objetivos principales del estudio son: (1) explorar el contenido emocional que capta el Termómetro de distrés (DT) en una muestra amplia de población española diagnosticada de cáncer usando el Inventario Breve de Síntomas-18 (BSI-18) y la categoría emocional de la Lista de Problemas (PL); y (2) examinar la precisión del DT como una herramienta de <i>screening</i> de malestar emocional usando como criterio estándar el BSI-18.</p> <p><b>Localización:</b> Fundación Instituto Valenciano de Oncología (FIVO).</p> <p><b>Periodo de realización:</b> Desde octubre de 2010 a octubre de 2012.</p>	<p><b>Población:</b> La población objeto de estudio son pacientes adultos oncológicos españoles. Los criterios de inclusión fueron: a) tener más de 18 años; b) tener un diagnóstico de cáncer; y c) ser capaz de proporcionar el consentimiento informado.</p> <p><b>Prueba a estudio:</b> El Termómetro de malestar (DT) es una medida de malestar emocional compuesta por un solo ítem, de carácter autoinformado. Se dan instrucciones a la persona para que indique en una escala analógica visual de 11 puntos cuánto malestar ha sentido en la semana anterior, desde 0 (Ausencia de malestar) a 10 (Malestar extremo). La definición de malestar de la NCCN® se presentó antes de completar el DT, para brindar a los y las pacientes un marco de referencia.</p> <p><b>Prueba de comparación:</b> La prueba utilizada como comparación es el Inventario Breve de Síntomas-18 (BSI-18), una lista de síntomas autoinformada. Los 18 ítems se miden en una escala tipo Likert de 5 puntos, que reflejan su malestar durante la semana anterior. Proporciona 3 puntuaciones de síntomas (ansiedad, depresión y somatización) y una puntuación total (Índice Global de Severidad). Usando los datos estandarizados por sexo que propuso Derogatis (2000), las puntuaciones se transforman en puntuaciones T para identificar los casos de malestar (<math>T \geq 63</math> en la puntuación total o en, al menos, dos subescalas). Estudios previos confirman sus propiedades psicométricas adecuadas.</p>	<p><b>Número de participantes:</b> 962 participantes que acuden a la FIVO para recibir tratamiento en cualquiera de sus departamentos médicos. Edad media = 57,55 años (<math>dt = 13,34</math>); 68 % mujeres. La mayoría de las y los participantes vivían con una pareja (72 %) y habían completado los estudios primarios (87 %). El tipo de cáncer primario más frecuente fue el de mama (41 %), seguido del ginecológico (16 %) y el de próstata (11 %). Sobre el 50 % de los y las participantes estaba en un estadio no avanzado de cáncer y el 60 % estaba recibiendo tratamiento médico en el momento del estudio.</p> <p><b>Criterios prueba diagnóstica positiva:</b> Este estudio explora el punto de corte óptimo del DT, por lo que en un principio no está predeterminado ni descrito.</p> <p><b>Criterios prueba de comparación positiva:</b> Con respecto a la puntuación de la escala BSI-18, las puntuaciones directas se transforman en puntuaciones T, usando los datos estandarizados por sexo propuestos por Derogatis (2000): <math>T \geq 63</math> en la puntuación total o en, al menos, dos subescalas para identificar casos de malestar.</p>	<p><b>Validez convergente:</b> La puntuación del DT estaba significativamente asociada con el GSI (<math>r = 0,65</math>) y las subescalas del BSI-18 (<math>r</math> ansiedad = 0,61; <math>r</math> depresión = 0,58; <math>r</math> somatización = 0,41). También se encontraron correlaciones positivas estadísticamente significativas entre el DT y los problemas emocionales recogidos en la Lista de Problemas: tristeza (0,53); preocupación (0,47); miedos (0,42); depresión (0,41) nerviosismo (0,41), y pérdida de interés (0,35).</p> <p><b>Precisión diagnóstica:</b> El área bajo la curva fue 0,82 (IC 95 % = 0,79 - 0,85), lo que indica que el DT tiene una buena precisión general para detectar casos de malestar acorde al BSI-18.</p> <p>El punto de corte óptimo estimado fue una puntuación de mayor o igual a 5 (sensibilidad = 0,90; especificidad = 0,64; VPP = 0,35; VPN = 0,97; CUI+ = 0,32; CUI- = 0,62). El valor global de la prueba muestra que el porcentaje de pacientes correctamente clasificados era del 68 %.</p>	<p>Las autoras exponen, como medida de la validez convergente del DT, sus correlaciones significativas con el BSI-18.</p> <p>Las autoras concluyen que el DT es una medida satisfactoria de malestar porque capta adecuadamente el rango de respuestas emocionales que la palabra malestar engloba. Consideran que la presentación de la definición de malestar previa a complementar el DT proporciona un marco de referencia desde el que los y las pacientes pueden evaluar su respuesta emocional, dejando de lado sus conceptualizaciones personales. Por tanto, debería contemplarse la inclusión de una definición de malestar con el DT, al menos en poblaciones de habla no inglesa.</p> <p>Concluyen una precisión general del DT satisfactoria como medida de <i>screening</i> con relación a una medida estandarizada de malestar psicológico (el BSI-18). Los índices de sensibilidad, especificidad, VPP, VPN, CUI+ y CUI- indican que el DT es una herramienta</p>	No genera preocupaciones respecto al riesgo de sesgo ni a la aplicabilidad.

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
Martínez, 2013 (97)		<p><b>Resultados analizados:</b> Se calculan las correlaciones (rho de Spearman) entre el DT y el BSI-18 (ítems, subescalas y GSI), así como entre el DT y los síntomas emocionales de la Lista de Problemas.</p> <p>Se calculan diferentes índices para estudiar la precisión diagnóstica y determinar el punto de corte óptimo: medidas globales (área bajo la curva ROC, y valor global de la prueba o proporción de casos correctamente clasificados), medidas de ocurrencia (sensibilidad y especificidad), medidas de discriminación (VPP y VPN) y medidas combinadas de ocurrencia y discriminación (utilidad clínica positiva (CUI+) y utilidad clínica negativa (CUI-).</p>			<p>adecuada para llevar a cabo un <i>screening</i> rutinario. Reflexionan que la clasificación errónea de ciertos falsos positivos pueda deberse a una malinterpretación en la escala de medida, que no sería interpretada como un continuo que va desde la «Ausencia de malestar» al «Malestar extremo», sino como una escala bidimensional afectiva que va desde el «Afecto positivo» al «Afecto negativo», donde una puntuación de 5 sería considerada como una condición emocionalmente neutral (ni bienestar, ni malestar). En este sentido, recomiendan un anclaje de punto medio próximo a la puntuación 5 etiquetado como «Malestar moderado» para reducir errores asociados a una posible malinterpretación de la dimensión afectiva del DT. Consideran que el desempeño del DT podría beneficiarse de la inclusión de otros instrumentos no-emocionales que ayudarían a reducir los falsos positivos, como el Termómetro de impacto o la Escala de Quejas Coloreada.</p>	

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
<b>Priede, 2014 (91)</b>	<p><b>Objetivos:</b> Evaluar las propiedades del Termómetro de distrés (DT) como prueba de cribado de trastornos ansioso-depresivos en una muestra española de población oncológica recién diagnosticada, utilizando dos pruebas de criterio distintas: un cuestionario autoaplicado de ansiedad y depresión (HADS) y una entrevista semiestructurada.</p> <p><b>Localización:</b> Hospital Universitario Marqués de Valdecilla (Santander, España).</p> <p><b>Periodo de realización:</b> De diciembre de 2011 a septiembre de 2013.</p>	<p><b>Población:</b> La población objeto de estudio son pacientes en edad adulta con un primer diagnóstico reciente de cáncer. Se han incluido pacientes con un primer diagnóstico confirmado por estudio de anatomía patológica. Los criterios de exclusión fueron: a) que hubiera pasado un tiempo desde el diagnóstico mayor a 4 meses; b) ser menor de 18 años; c) tener más de 70 años; d) ser un o una paciente con recidiva o que haya padecido otro proceso oncológico con anterioridad; e) tener discapacidad intelectual, deterioro de las funciones cognitivas o incapacidad para comprender las pruebas de evaluación.</p> <p><b>Prueba a estudio:</b> La prueba índice es la prueba del Termómetro de distrés (DT), un cuestionario de un solo ítem que evalúa la presencia de malestar a través de una escala de 11 puntos en la que la puntuación 0 significa ausencia de malestar y la puntuación 10 hace referencia a un distrés máximo. El DT se aplica de manera heteroadministrada durante la entrevista diagnóstica.</p>	<p><b>Número de participantes:</b> La muestra inicial estaba compuesta por 215 pacientes que cumplían los criterios de inclusión en el estudio. La muestra final seleccionada estuvo compuesta por un total de 135 participantes. Los grupos de participantes frente a los no participantes solo presentaron diferencias significativas en nivel educativo (superior entre las y los participantes) y en antecedentes psicopatológicos (superior entre las y los participantes). El nivel educativo no influyó en los niveles de malestar emocional, pero los antecedentes psicopatológicos sí se relacionaron con la presencia de malestar emocional. Del total de 135 participantes, el 65,19 % eran mujeres, y la edad media fue de 52,84 años (dt = 10,15). El 76,30 % estaba casado o casada o en pareja. El diagnóstico oncológico más frecuente fue el cáncer de mama (33,33 %) y de pulmón (14,81 %).</p> <p><b>Criterios prueba diagnóstica positiva:</b> Este estudio explora los puntos de corte óptimos del DT, por lo que en un principio no están predeterminados ni descritos.</p>	<p><b>Validez convergente:</b> La correlación entre el DT y la HADS fue <math>r = 0,64</math> (<math>p &lt; 0,0005</math>); entre el DT y la subescala de ansiedad de la HADS fue de <math>r = 0,66</math> (<math>p &lt; 0,0005</math>) y con la subescala de depresión fue de <math>r = 0,51</math> (<math>p &lt; 0,0005</math>).</p> <p><b>Precisión diagnóstica:</b> El área bajo la curva para el DT usando como referencia la HADS fue 0,79 (IC 95 % = 0,71 - 0,86; <math>p &lt; 0,0005</math>). De acuerdo con el índice de Youden, el punto de corte óptimo queda establecido en 4 (<math>J = 0,45</math>). Para dicho punto, su sensibilidad es 0,89, su especificidad = 0,56, el VPP = 0,43 y el VPN = 0,93.</p> <p>El área bajo la curva del DT usando como referencia la medida combinada (HADS y entrevista clínica) fue de 0,86 (IC 95 % = 0,79 - 0,93; <math>p &lt; 0,0005</math>), siendo su punto de corte óptimo 6 (<math>J = 0,58</math>). Para este punto, la sensibilidad fue 0,82, la especificidad = 0,76, el VPP = 0,40 y el VPN = 0,95.</p>	<p>El equipo autor del estudio concluye que el DT es una prueba de cribado adecuada para detectar el malestar emocional en pacientes a quienes han diagnosticado recientemente de cáncer.</p> <p>Cuando se emplea la HADS como prueba de referencia, su capacidad discriminativa es, al menos, aceptable para un punto de corte de 4 puntos.</p> <p>Al emplear como prueba de referencia la entrevista clínica combinada con la HADS, la capacidad discriminativa del DT aumenta para un punto de corte de 6, considerándose buena.</p>	<p>La selección de pacientes presenta riesgo de sesgo incierto, así como preocupaciones sobre su aplicabilidad.</p> <p>También existe riesgo de sesgo en la prueba de referencia y en el flujo y tiempos.</p> <p>No genera preocupaciones sobre la aplicabilidad de las pruebas.</p>

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
Priede, 2014 (91)		<p><b>Prueba de comparación:</b> La prueba de comparación es la Escala Hospitalaria de Ansiedad y Depresión (HADS). Consta de 14 ítems, divididos en dos subescalas de depresión y ansiedad. La consistencia interna de la escala global es alta (<math>\alpha = 0,90</math>), así como la de sus subescalas (<math>\alpha</math> ansiedad = 0,85; <math>\alpha</math> depresión = 0,84). El o la paciente responde esta escala de manera autoaplicada en su domicilio y después lo entrega en el hospital. Por otra parte, dado que la HADS es a su vez una prueba de cribado, se realizó un segundo análisis en el que se combinó la clasificación realizada por medio de la HADS con una entrevista diagnóstica semiestructurada basada en los criterios diagnósticos DSM-IV.</p> <p><b>Resultados analizados:</b> Primero, se empleó como prueba de referencia la clasificación aportada por la HADS y, después, la combinación de la HADS con la clasificación realizada por la entrevista diagnóstica. Se realizó un análisis de la curva ROC para determinar el punto de corte óptimo para identificar a los y las pacientes que presentan malestar emocional significativo (tanto para la HADS como para la medida combinada). Se generaron las curvas ROC y se calculó la sensibilidad, la especificidad y el área bajo la curva. La determinación del punto de corte óptimo se efectuó maximizando el índice de Youden (<math>J = \text{sensibilidad} + \text{especificidad} - 1</math>). También se calculó el VPP y el VPN.</p>	<p><b>Criterios prueba de comparación positiva:</b> Con respecto a la prueba HADS, se empleó la puntuación de 15 como punto de corte para discriminar los casos. Cuando se utilizó como prueba de comparación la combinación de la HADS con una entrevista semiestructurada basada en los criterios diagnósticos DSM-IV, se consideró como casos solamente a las personas pacientes detectadas de modo convergente por ambas pruebas, es decir, aquellas con una puntuación mayor o igual a 15 en la HADS que también fuesen consideradas como casos mediante entrevista clínica.</p>			

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
<b>Martínez, 2015 (98)</b>	<p><b>Objetivos:</b> Investigar los efectos de añadir el IT al DT comparando la precisión diagnóstica de los termómetros tanto de manera individual como combinada (puntuaciones pareadas y puntuaciones sumadas) en población oncológica adulta.</p> <p><b>Localización:</b> Fundación Instituto Valenciano de Oncología (FIVO, Valencia).</p> <p><b>Periodo de realización:</b> De mayo de 2013 a febrero de 2014.</p>	<p><b>Población:</b> La población objeto de estudio son pacientes con cáncer en España. Los criterios de inclusión fueron: a) ser mayor de 18 años; b) tener un diagnóstico de cáncer; y c) ser capaz de proporcionar el consentimiento informado.</p> <p><b>Prueba a estudio:</b> La prueba de interés es el Termómetro de distrés (DT). El DT es una medida autoinformada, compuesta por un único ítem que emplea una escala analógica visual de 11 puntos —desde 0 (Ningún malestar emocional) a 10 (Malestar emocional extremo)— para evaluar el malestar emocional experimentado durante la última semana. La versión en español de esta herramienta ha demostrado buena precisión diagnóstica. También se empleó el DT en combinación con el Termómetro de impacto (IT) como prueba de interés (tanto con puntuaciones pareadas como puntuaciones sumadas). El IT es una medida autoinformada, compuesta por un único ítem que emplea una escala analógica visual de 11 puntos —desde 0 (Ausencia de interferencia) a 10 (Interferencia extrema)— para determinar la interferencia que el malestar emocional ha tenido en su vida cotidiana.</p>	<p><b>Número de participantes:</b> Se invitó a participar a una muestra consecutiva de pacientes que visitaban algún departamento médico en la Fundación Instituto Valenciano de Oncología (FIVO). El 6 % de las y los pacientes que se abordaron (25 de 410) declinaron participar. Las razones esgrimidas fueron: no tener interés (45 %), estar demasiado cansados (28 %), tener mucha prisa (10 %) y otras razones (17 %). La muestra final estuvo compuesta por 385 pacientes oncológicos. Edad media = 58,7 años (dt = 13,2). El 55 % eran hombres. La mayoría de las y los participantes convive con su pareja (73 %) y ha completado la escuela primaria (89 %). El tipo de cáncer primario más frecuente fue el cáncer de mama (24 %), seguido por el cáncer de próstata (20 %) y el melanoma (15 %). Más de la mitad de los y las participantes estaban en un estadio no avanzado de cáncer, y aproximadamente el 62 % estaba recibiendo tratamiento médico en el momento del estudio.</p>	<p><b>Validez convergente:</b> Respecto a la validez convergente del DT y el IT, las correlaciones con el BSI18 fueron significativas (<math>r</math> DT = 0,62; <math>r</math> DIT = 0,63; <math>r</math> IT = 0,56).</p> <p><b>Precisión diagnóstica:</b> Para el DT, el valor de área bajo la curva (AUC) fue = 0,84 (IC 95% = 0,80 - 0,87). El punto de corte óptimo estimado fue mayor o igual a 5 (sensibilidad = 0,93; especificidad = 0,67; VPP = 0,25; VPN = 0,99; CUI+ = 0,23; CUI- = 0,66).</p> <p>Para el DIT (suma de puntuaciones del DT y el IT), el valor de área bajo la curva (AUC) fue = 0,83 (95 % IC = 0,79 - 0,87). El punto de corte óptimo estimado fue mayor o igual a 9 (sensibilidad = 0,83; especificidad = 0,71; VPP = 0,26; VPN = 0,97; CUI+ = 0,22; CUI- = 0,69).</p> <p>Para la combinación pareada del DT y el IT, el punto de corte óptimo estimado fue mayor o igual a 5 para ambas escalas (sensibilidad = 0,74; especificidad = 0,79; VPP = 0,30; VPN = 0,96; CUI+ = 0,22; CUI- = 0,76).</p> <p>Las dos combinaciones de DT e IT alcanzan una mayor especificidad frente al DT solo. Sin embargo, ninguna de las dos combinaciones proporciona un índice significativamente mejor que los obtenidos por el DT.</p>	<p>Las autoras concluyen que el DT es una buena herramienta para el <i>screening</i> de malestar emocional en pacientes con cáncer en España, mostrando una precisión diagnóstica adecuada comparada con el BSI18 y una validez convergente satisfactoria. En base a sus resultados, las autoras concluyen que añadir el IT al DT (ya sea como suma de puntuaciones o puntuaciones pareadas) no mejora significativamente la precisión diagnóstica del DT como herramienta de <i>screening</i> para el malestar emocional. No obstante, también apuntan que dicho uso combinado puede ser interesante para el personal de la salud, pues recopila más información relevante sobre la respuesta emocional de los y las pacientes sin cambiar el proceso de <i>screening</i> (tiempo y puntuación). Su uso queda, entonces, a discreción de los/las profesionales de la salud.</p>	<p>No genera preocupaciones respecto al riesgo de sesgo ni a la aplicabilidad.</p>

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
Martínez, 2015 (98)		<p><b>Prueba de comparación:</b> La prueba de comparación es el Inventario Breve de Síntomas (BSI-18). Se trata de un autoinforme compuesto por 18 ítems (escala Likert de 5 puntos) que evalúa el malestar emocional durante la semana anterior. El instrumento ofrece puntuaciones en somatización, depresión y ansiedad y una puntuación total (el Índice Global de Severidad -GSI-) en malestar emocional. El BSI-18 ha mostrado propiedades psicométricas adecuadas en estudios con población oncológica española y ha sido empleado como prueba de comparación del DT.</p> <p><b>Resultados analizados:</b> Se calcularon correlaciones entre el BSI-18 y las medidas ultracortas (DT e IT), mediante la rho de Spearman, para estimar la validez convergente del DT e IT (individualmente y en combinación). Se calcularon varios índices para estudiar la precisión diagnóstica y determinar el punto de corte óptimo: el área bajo la curva (AUC) y las diferencias estadísticas entre las curvas ROC; la sensibilidad y la especificidad (medidas de ocurrencia); el VPP y el VPN (medidas de discriminación); la CUI+ y la CUI- (medidas combinadas de ocurrencia y discriminación).</p>	<p><b>Criterios prueba diagnóstica positiva:</b> Este estudio explora la precisión diagnóstica del DT, y del DT en combinación con el IT, por lo que los puntos de corte no están predeterminados y descritos en un principio.</p> <p><b>Criterios prueba de comparación positiva:</b> Para el BSI-18, las puntuaciones se transforman en puntuaciones T para identificar casos de malestar, utilizando los datos normatizados por sexo propuestos por Derogatis (<math>T \geq 63</math> en el GSI o, al menos, en dos subescalas).</p>			

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
<b>Andreu, 2015 (35)</b>	<p><b>Objetivos:</b> Extender la investigación previa sobre la utilidad y la precisión diagnóstica del Termómetro de distrés (DT), así como sobre sus posibilidades de mejora —incorporando el IT— en un grupo de pacientes españoles adultos con neoplasias hematológicas.</p> <p><b>Localización:</b> Hospital Universitario la Fe de Valencia.</p> <p><b>Periodo de realización:</b> De febrero de 2013 a febrero de 2014.</p>	<p><b>Población:</b> La población objeto de estudio son pacientes adultos oncológicos hospitalizados con neoplasias hematológicas en España. Los criterios de elección fueron: a) tener un diagnóstico de neoplasia hematológica; b) no encontrarse en situación de aislamiento; y c) tener 18 años o más.</p> <p><b>Prueba a estudio:</b> La prueba de interés es el Termómetro de distrés (DT). El DT emplea una escala visual de 11 puntos —desde 0 (Ningún malestar emocional) a 10 (Malestar emocional extremo)— para evaluar el malestar emocional experimentado durante la última semana. También se utilizó el DT en combinación con el Termómetro de impacto (IT) como prueba de interés (tanto con puntuaciones pareadas como puntuaciones sumadas). El IT emplea una escala visual de 11 puntos —desde 0 (Ausencia de interferencia) a 10 (Interferencia extrema)— para determinar la interferencia en la vida cotidiana del malestar emocional.</p> <p><b>Prueba de comparación:</b> La prueba de comparación es el Inventario Breve de Síntomas (BSI-18). Se trata de un autoinforme compuesto por 18 ítems (escala Likert de 5 puntos) que evalúa el malestar emocional. El instrumento ofrece puntuaciones en somatización, depresión y ansiedad y una puntuación total (GSI) en malestar emocional. El BSI-18 ha mostrado propiedades psicométricas satisfactorias en población oncológica española.</p> <p><b>Resultados analizados:</b> Se calcularon correlaciones bivariadas (prueba rho no paramétrica de Spearman) para explorar la validez convergente de los termómetros contemplados (DT, IT, DIT). La precisión diagnóstica se exploró a través de varios índices: área bajo la curva (AUC), medidas de ocurrencia (sensibilidad y especificidad) y medidas de discriminación (VPP y VPN). También se calcularon los índices de utilidad clínica positiva y negativa.</p>	<p><b>Número de participantes:</b> La muestra inicial la formaron 100 pacientes consecutivos ingresados en el Servicio de Hematología del Hospital Universitario la Fe de Valencia. Un 19 % rechazó participar en el estudio, por lo que la muestra final estaba compuesta por 81 pacientes. Alrededor del 60 % tiene entre 40 y 60 años (M = 51 años; dt = 10,7) y estaba laboralmente activo o activa antes del ingreso hospitalario. Un 53 % son varones, un 80 % tiene pareja y un 96 % cursó al menos estudios primarios. Todas y todos los pacientes están ingresados en el hospital debido al propio tratamiento (90 %) o a complicaciones médicas (10 %).</p> <p><b>Criterios prueba diagnóstica positiva:</b> Este estudio explora los puntos de corte óptimos del DT y del DT en combinación con el IT, por lo que en un principio no están definidos.</p> <p><b>Criterios prueba de comparación positiva:</b> Se consideran casos con malestar clínicamente significativo a aquellos con una puntuación <math>T \geq 63</math> en la puntuación total o, al menos, en dos de las subescalas.</p>	<p><b>Validez convergente:</b> Los tres termómetros correlacionaron significativamente con las tres subescalas del BSI-18 y la puntuación total (GSI), siendo las correlaciones más altas las obtenidas por el DT (<math>r</math> GSI = 0,58; <math>r</math> ansiedad = 0,57; <math>r</math> depresión = 0,54; <math>r</math> somatización = 0,38).</p> <p><b>Precisión diagnóstica:</b> Para el DT, el valor de área bajo la curva (AUC) resultó ser satisfactorio (AUC = 0,82; IC 95 % = 0,73 - 0,91). El punto de corte óptimo estimado fue mayor o igual a 5 (sensibilidad = 0,94; especificidad = 0,48; VPP = 0,42; VPN = 0,93; CUI+ = 0,39; CUI- = 0,45).</p> <p>Para el DIT (suma de puntuaciones del DT y el IT), el valor de área bajo la curva (AUC) fue = 0,77 (IC 95 % = 0,64 - 0,88). El punto de corte óptimo estimado fue mayor o igual a 9 (sensibilidad = 0,82; especificidad = 0,62; VPP = 0,35; VPN = 0,90; CUI+ = 0,29; CUI- = 0,56).</p> <p>Para la combinación pareada del DT y el IT, el punto de corte óptimo estimado fue mayor o igual a 5 para ambas escalas (sensibilidad = 0,53; especificidad = 0,91; VPP = 0,60; VPN = 0,88; CUI+ = 0,32; CUI- = 0,80).</p> <p>Ninguna de las dos combinaciones del DT y el IT mostró resultados significativamente mejores que los obtenidos por el DT.</p>	<p>Las autoras concluyen un buen comportamiento del DT como instrumento de cribado de distrés.</p> <p>Su validez convergente queda avalada por la correlación significativa con la puntuación total (GSI) y con cada una de las subescalas del BSI-18.</p> <p>Los resultados respaldan la precisión diagnóstica del DT como herramienta de cribado. Con respecto a la combinación mediante suma del DT con el IT, la exactitud diagnóstica no fue significativamente diferente de la correspondiente al DT. Por otra parte, cuando el DT y el IT se combinaron con puntuaciones pareadas, la sensibilidad experimentó una caída drástica (situándola en niveles inaceptables para una herramienta de cribado).</p> <p>Las autoras concluyen que la inclusión del IT no mejora de manera sustantiva la precisión del DT como herramienta de cribado de distrés.</p>	No genera preocupaciones respecto al riesgo de sesgo ni a la aplicabilidad.

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
<b>Andreu, 2019 (99)</b>	<p><b>Objetivos:</b> Examinar la precisión del DT para detectar malestar clínicamente significativo según lo evaluado por el BSI-18 (medida criterio) en personas españolas supervivientes de cáncer de inicio en la edad adulta postratamiento.</p> <p><b>Localización:</b> Fundación Instituto Valenciano de Oncología (FIVO, Valencia).</p> <p><b>Periodo de realización:</b> No se explicita.</p>	<p><b>Población:</b> La población objeto de estudio son personas españolas supervivientes tras tratamiento de cáncer de inicio en la edad adulta. Los criterios de inclusión fueron: a) tener 18 años o más; b) haber tenido un diagnóstico de cáncer, pero sin síntomas o signos de cáncer en ese momento; c) haber finalizado la cirugía, la quimioterapia o la radioterapia al menos 4 semanas antes del estudio; d) tener conocimientos de lengua española.</p> <p><b>Prueba a estudio:</b> La prueba diagnóstica de interés es el Termómetro de distrés (DT), una escala analógica visual de un solo ítem que mide el malestar emocional. Los individuos indican en una escala de 11 puntos desde 0 (Ausencia de malestar) a 10 (Malestar extremo)— cuánto malestar han sentido en los siete días anteriores.</p> <p><b>Prueba de comparación:</b> La prueba de comparación es el Inventario Breve de Síntomas (BSI-18; Derogatis, 2013), una lista de verificación de síntomas de autoinforme que comprende 18 ítems calificados en una escala tipo Likert de 5 puntos. Los individuos han de valorar cada ítem según cómo se han sentido durante los últimos siete días. La escala proporciona una medida global de malestar emocional (el Índice Global de Severidad, GSI) y tres síntomas nucleares (somatización, depresión y ansiedad). El BSI-18 ha mostrado propiedades psicométricas satisfactorias en estudios en población oncológica española.</p>	<p><b>Número de participantes:</b> Se invitó a participar a una muestra consecutiva inicial de 254 participantes que acudían a la FIVO. El 93 % (236) dieron su consentimiento informado y completaron el conjunto de cuestionarios. La muestra final está formada por 236 participantes. Edad media = 58 años (dt = 13,4). 65 % mujeres. Más de dos tercios de los y las participantes conviven con su pareja (69 %), la mayoría ha completado la educación primaria (89 %), el 43 % estaba en la jubilación o de baja laboral, y el 11 %, sin empleo. El tipo de cáncer más frecuente fue el cáncer ginecológico (36 %), seguido del cáncer de próstata (18 %), el cáncer de mama (16 %) y el melanoma (13 %). Casi una de cada cuatro personas participantes (24 %) había completado el tratamiento primario hacía menos de un año, el 43 % había finalizado el tratamiento entre 1 y 5 años antes, y el 33 % llevaba más de 5 años postratamiento.</p> <p><b>Criterios prueba diagnóstica positiva:</b> Dado que el objetivo del estudio es examinar la precisión del DT para detectar malestar clínicamente significativo en personas españolas supervivientes de cáncer de inicio en la edad adulta postratamiento, no se define <i>a priori</i> el valor umbral o punto de corte.</p>	<p><b>Validez convergente:</b> Las puntuaciones del DT correlacionaron significativamente con las del BSI-18 (<math>r</math> GSI = 0,67; <math>r</math> somatización = 0,42; <math>r</math> depresión = 0,56; <math>r</math> ansiedad = 0,62).</p> <p><b>Precisión diagnóstica:</b> Se calculó el área bajo la curva (AUC = 0,81; IC 95 % = 0,74 - 0,87), que indica que el DT muestra una precisión global adecuada. Se consideró una puntuación de 5 como el punto de corte óptimo. El valor global de la prueba para este punto fue de 64 %, lo que significa que la prueba clasifica correctamente el 64 % de los y las pacientes. La sensibilidad fue del 0,89; la especificidad, 0,59; el VPP fue 0,29 y el VPN, 0,97. La CUI- fue de 0,57 y la CUI+ fue de 0,26.</p>	<p>Las autoras concluyen que el DT muestra una validez convergente satisfactoria, correlacionando significativamente tanto con la puntuación global del BSI-18, como con cada una de las subescalas. Además, el DT muestra una precisión global adecuada, teniendo en cuenta el BSI-18 como estándar oro.</p> <p>Para seleccionar el punto de corte óptimo, las autoras consideraron un criterio combinado: una sensibilidad óptima (para asegurar que los posibles casos de morbilidad psicológica no se pierdan) junto con índices de utilidad clínica que combinan ocurrencia y medidas de discriminación. En el contexto de supervivientes de cáncer, la sobredetección es preferible a la infradetección, debido al contacto intermitente con el personal de la salud y a que existen menos oportunidades para identificar el malestar emocional.</p>	No genera preocupaciones respecto al riesgo de sesgo ni a la aplicabilidad.

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
Andreu, 2019 (99)		<p><b>Resultados analizados:</b> Para analizar la validez convergente del DT, se analizaron las correlaciones entre el DT y el BSI-18 (GSI y subescalas), calculando la rho de Spearman.</p> <p>Para estudiar la precisión diagnóstica del DT y determinar el punto de corte óptimo para el malestar clínicamente significativo, se calcularon el área bajo la curva (AUC), el valor global de la prueba, la sensibilidad, la especificidad, el VPP, el VPN, la CUI+ y la CUI-.</p>	<p><b>Criterios prueba de comparación positiva:</b> Con respecto al BSI-18, las puntuaciones se transforman en puntuaciones T para identificar casos de malestar clínicamente significativo usando datos normativos específicos por sexo (las puntuaciones <math>T \geq 63</math> en el GSI o en, al menos, dos subescalas se clasifican como caso)</p>		<p>Los resultados sugieren que el DT es más adecuado para un <i>screening</i> rutinario que para la búsqueda de casos, dado el gran número de falsos positivos. En base a los resultados obtenidos y en consonancia con las directrices de la NCCN®, en la fase de supervivencia del cáncer, se recomienda que el uso del DT se haga en el marco de un <i>screening</i> de dos pasos, en el que el DT sirva como primer paso para identificar no-casos y, posteriormente, se añada una segunda herramienta de <i>screening</i> con más ítems para reducir los falsos positivos y evitar la sobrecarga de los escasos recursos de salud disponibles.</p>	

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
<b>Limonero, 2012 (102)</b>	<p><b>Objetivos:</b> Analizar las propiedades psicométricas del Cuestionario de Detección de Malestar Emocional (DME) y confirmar su posible validez clínica.</p> <p><b>Localización:</b> Estudio multicéntrico en unidades de cuidados paliativos en cinco hospitales de Cataluña (Consorti Sanitari de Terrassa, Institut Català d'Oncologia, Corporació Sanitària Parc Taulí de Sabadell, Hospital Universitari Sant Joan de Reus y la Fundació Hospital-Residència Sant Camil de Sant Pere de Ribes).</p> <p><b>Periodo de realización:</b> De marzo a junio de 2010.</p>	<p><b>Población:</b> La población objeto de estudio es la población oncológica-paliativa. Los criterios de inclusión fueron: a) pacientes con enfermedad oncológica avanzada hospitalizados; b) mayores de 18 años; c) consentimiento informado. Se excluyeron los y las pacientes que no cumplían estos criterios o que, aun cumpliéndolos, estuvieran en una situación clínica que impidiera responder el cuestionario, presentaran deterioro cognitivo según la escala Pfeiffer (<math>\geq 2</math> errores) o no fueran capaces de mantener una conversación o comprender las preguntas del cuestionario.</p> <p><b>Prueba a estudio:</b> La prueba de interés es el Cuestionario de Detección de Malestar Emocional (DME) en pacientes oncológicos al final de la vida. Consta de dos partes. La primera contiene tres preguntas dirigidas a la persona enferma, dos de ellas en formato de escala visual numérica de 0 a 10, en las que se evalúa el estado de ánimo y la percepción de afrontamiento de la situación, y otra que registra la presencia o la ausencia de preocupaciones. La segunda parte consiste en una observación, por parte de la o el profesional de la salud, acerca de la presencia de signos externos de malestar emocional, en donde se registra su existencia. El DME otorga una puntuación total (0 a 20) formada por la suma de las puntuaciones de las respuestas a las preguntas referentes a la valoración del estado de ánimo y a la de cómo lleva la situación. Las preguntas sobre preocupaciones o signos externos de malestar emocional permiten al personal sanitario llevar a cabo una atención más específica, pero no se valoran cuantitativamente.</p>	<p><b>Número de participantes:</b> Del total de pacientes consecutivos analizados, se excluyeron cuatro (3,6 %) por no cumplir alguno de los requisitos, de forma que la muestra final quedó constituida por 105 pacientes (50 hombres y 55 mujeres) con una media de edad de <math>67,3 \pm 12,6</math> años. El número de pacientes recogido por cada centro fue de 25, excepto en uno de ellos (la Fundació Hospital-Residència Sant Camil de Sant Pere de Ribes), donde por cuestiones organizativas se registraron datos de 9 pacientes. Los diagnósticos principales más comunes fueron: cáncer de pulmón o pleura (16,2 %), cáncer de colon, recto y ciego (15,2 %), páncreas (13,3 %) y mama (11,4 %).</p> <p><b>Criterios prueba diagnóstica positiva:</b> Dado que el objetivo del estudio es analizar las propiedades psicométricas del DME, el punto de corte adecuado de la escala no se establece <i>a priori</i>.</p> <p><b>Criterios prueba de comparación positiva:</b> El punto de corte establecido para la entrevista clínica psicológica fue mayor o igual a 3. En el caso del DT, se considera que un o una paciente presenta malestar si tiene una puntuación mayor o igual a 4. Con respecto a la HADS, la suma de ambas escalas proporciona una medida de malestar global, con un punto de corte recomendado mayor o igual a 16.</p>	<p><b>Validez convergente:</b> El DME correlaciona positiva y significativamente con las puntuaciones del DT (<math>r = 0,235</math>), la valoración psicológica clínica (<math>r = 0,438</math>), la HADS (0,513) y sus subescalas de ansiedad (<math>r = 0,614</math>) y depresión (<math>r = 0,499</math>).</p> <p><b>Consistencia interna:</b> El alfa de Cronbach fue de 0,685.</p> <p><b>Precisión diagnóstica:</b> Se analizaron las curvas ROC para determinar el punto de corte óptimo del DME. Utilizando la valoración clínica psicológica como prueba de referencia, se observa un área bajo la curva de 0,81 (IC 95 % = 0,72 - 0,9), para un punto de corte mayor o igual a 9,25. Si se lleva a cabo el mismo análisis tomando como variable de comparación la puntuación obtenida en el DT, se observa un área bajo la curva de 0,85 (IC 95 % = 0,77 - 0,94), para un punto de corte mayor o igual a 9,75. Si se coteja la puntuación global del DME con la escala HADS, se observa un área bajo la curva de 0,92 (IC 95 % = 0,84 - 0,99), para un punto de corte de 9,25. Teniendo en cuenta esto, se establece que, para el DME, la puntuación global de corte será mayor o igual a 9 (no queda claro por qué optan por ese punto de corte en vez del 10). También se calcula el VPP y el VPN para el punto de corte 9, en comparación con la valoración clínica (VPP = 0,80; VPN = 0,782), el DT (VPP = 0,808; VPN = 0,792), y la HADS (VPP = 1; VPN = 0,687).</p>	<p>El equipo autor del estudio concluye la validez clínica del DME como herramienta de cribado de malestar emocional en las personas enfermas oncológicas avanzadas hospitalizadas en unidades de cuidados paliativos, de forma fácil, rápida, ética y no invasiva.</p>	<p>Riesgo de sesgo con respecto a la prueba índice.</p> <p>No genera preocupaciones en el resto de los dominios ni en la aplicabilidad de los resultados.</p>

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
<p><b>Limonero, 2012 (102)</b></p>		<p><b>Prueba de comparación:</b>            Para evaluar la validez convergente de la prueba de interés, se emplearon dos herramientas. Por una parte, el Termómetro de distrés (DT), que es una escala visual numérica de un solo ítem que puntúa de 0 a 10 y mide la valoración global de malestar emocional que experimenta la o el paciente. Por otra parte, se empleó también la Escala Hospitalaria de Ansiedad y Depresión (HADS). Consta de 14 ítems con formato de respuesta tipo Likert de 4 puntos (rango de 0 a 3); siete corresponden a la subescala de ansiedad y los otros siete a la de depresión. La puntuación para cada subescala es de 21 puntos</p> <p>Para evaluar la precisión diagnóstica de la prueba de interés, se empleó como prueba de comparación una entrevista clínica semiestructurada realizada de forma independiente por un psicólogo con experiencia en cuidados paliativos. A través de ella, se evaluaba globalmente el malestar emocional de la persona enferma con una escala tipo Likert de cinco puntos, en la cual 1 indica «Ningún malestar» y 5 indica «Malestar emocional máximo». La duración de la entrevista oscilaba entre 15 y 20 minutos.</p> <p>Aunque en el apartado de método no se especifica como tal, el DT y la HADS también se utilizan como medidas criterio para evaluar la precisión diagnóstica del DME.</p> <p><b>Resultados analizados:</b>            Se analizaron los coeficientes de correlación intraclase del cuestionario DME con el DT, la valoración clínica psicológica, la HADS y sus subescalas.</p> <p>También se calculó el alfa de Cronbach como medida de consistencia interna de la escala.</p> <p>Para la precisión diagnóstica, se realizó un análisis de la curva ROC y se calcularon la sensibilidad, la especificidad, el VPP y el VPN.</p>				

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
<b>Mateo-Ortega 2019 (103)</b>	<p><b>Objetivos:</b> Desarrollar un instrumento para evaluar mejor las necesidades psicosociales y el malestar emocional en pacientes oncológicos al final de la vida.</p> <p><b>Localización:</b> Unidades de cuidados paliativos de 19 hospitales participantes.</p> <p><b>Periodo de realización:</b> El reclutamiento de pacientes duró 3 meses (en 2017).</p>	<p><b>Población:</b> La población objeto de estudio son pacientes al final de la vida. Los criterios de inclusión son: (1) diagnóstico de cáncer avanzado o terminal; (2) edad mayor o igual a 18 años; (3) consentimiento informado; y (4) estar recibiendo tratamiento de los equipos psicosociales y de cuidados paliativos. Los criterios de exclusión son: (1) presencia de condiciones que impedirían al paciente completar el cuestionario; (2) deterioro cognitivo (puntuación del Pfeiffer test mayor a 3); o (3) negarse a dar el consentimiento informado.</p> <p><b>Prueba a estudio:</b> La prueba de interés es el Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida (ENP-E), dividido en tres secciones. La sección A consta de 13 ítems que preguntan al paciente sobre sentimientos, pensamientos y actividades que ha hecho recientemente. La escala de respuesta es tipo Likert de 5 puntos. La sección B consta de seis preguntas sobre el tipo de preocupaciones que tiene el o la paciente. La escala de respuesta es tipo Likert de 5 puntos. La sección C consta de seis preguntas sobre signos externos de malestar emocional detectados por quien entrevista.</p>	<p><b>Número de participantes:</b> La muestra seleccionada estuvo formada por 150 pacientes que reciben tratamiento en las unidades de cuidados paliativos de los 19 hospitales participantes y que se reclutaron durante los tres días siguientes al ingreso hospitalario: 21 pacientes formaron parte del estudio piloto y 129 pacientes formaron parte de la fase de estudio. La información de un paciente fue descartada por estar incompleta. El 51 % eran mujeres. La edad media en hombres fue de 72 años (rango entre 33 y 93); la edad media en mujeres fue de 73 años (rango entre 43 y 93). El 49 % de las personas participantes estaban casadas, el 16,8 % estaban solteras, el 7 % estaban divorciadas o separadas, y el 20 % estaban viudas. Todas tenían cáncer avanzado, principalmente de colon (15,4 %), de pulmón (15,4 %), pancreático (8,7 %), de mama (8,1 %), o de estómago. La mayoría tenía metástasis (69,8 %) o cáncer localizado (13,4 %). La mayoría de los y las pacientes (67,1 %) no tenía historial de abuso de tabaco, alcohol u otras drogas; el 16 % tenía historial de tabaquismo, y el 10,7 %, de alcoholismo. La adicción al cannabis u otras drogas era poco común (1,4 %). La mayoría de las y los pacientes no tenía historial de enfermedad psiquiátrica (79,9 %), el 12,8 % tenía historial de trastorno depresivo y el 2 % de ansiedad. Solo tres pacientes (2,1 %) tenían historial de trastornos adaptativos o psicóticos.</p>	<p><b>Estructura factorial</b> Análisis factorial exploratorio (AFE): índice KMO = 0,75; prueba de esfericidad de Bartlett (<math>p &lt; 0,01</math>). La solución propuesta fue cuatrifactorial (factor 1: bienestar emocional; factor 2: apoyo social; factor 3: espiritualidad; factor 4: información).</p> <p><b>Consistencia interna:</b> Alfa de Cronbach = 0,74 o 0,75 (aparecen ambos datos); alfa bienestar emocional = 0,77; alfa apoyo social = 0,64; alfa espiritualidad = 0,48. El factor 4 estaba formado por un único ítem, por lo que la consistencia interna no se calculó.</p> <p><b>Validez convergente:</b> ICC ENP-E – HADS = 0,67 ICC ENP-E – ítem 15 EORTC QLQ-C15 = -0,0344 ICC ENP-E – DT = 0,516 ICC ENP-E – Percepción del tiempo = 0,326.</p> <p><b>Precisión diagnóstica:</b> Área bajo la curva (IC 95 %) = 0,82 (0,767 - 0,897); punto de corte establecido <math>\geq 28</math>; sensibilidad = 0,746; especificidad = 0,744.</p> <p>Además, algunos ítems (1, 2, 4, 10 y 12) discriminan mejor que otros, por lo que fueron considerados preguntas clave, de manera que puntuaciones mayores o iguales a 4 en alguno de estos ítems se consideran posibles signos de alerta de necesidades psicosociales y espirituales.</p>	<p>El equipo autor concluye que el ENP-E es un cuestionario breve y sencillo para evaluar de manera comprensiva las necesidades psicológicas y espirituales. La consistencia interna fue adecuada, así como las demás propiedades psicométricas. Su valor clínico está apoyado por el tiempo relativamente corto para completarlo, junto con la claridad de las preguntas percibida por parte de los y las pacientes.</p>	<p>Riesgo de sesgo incierto con respecto a la selección de pacientes y la prueba de referencia. No genera preocupaciones sobre la aplicabilidad.</p>

Cita abreviada	Estudio	Pregunta de investigación	Método	Resultados	Conclusiones	Calidad del estudio
Mateo-Ortega 2019 (103)		<p><b>Prueba de comparación:</b> Emplean la Escala Hospitalaria de Ansiedad y Depresión (HADS), concretamente la versión española de Tejero <i>et al.</i> (1986); el ítem de calidad de vida del EORTC QLQC15-PAL; el Termómetro de distrés (Holland, 1999); la percepción subjetiva del paso del tiempo (Bayés <i>et al.</i>, 1995, 1997). La prueba utilizada como comparación fue una combinación de los valores de la HADS (HADS mayor o igual a 17), junto con la puntuación del ítem de calidad de vida del EORTC QLQ-C15-PAL (menor o igual a 3), la puntuación del DT (mayor a 5) y el grado de sufrimiento (percepción a largo plazo). La prueba de referencia se construyó para indicar si el valor es positivo o no (dicotómica).</p> <p><b>Resultados analizados:</b> Validez convergente: análisis del índice de correlación intraclass (ICC) entre la puntuación obtenida en el ENP-E y las puntuaciones obtenidas en las escalas validadas empleadas (HADS, QLQ-C15-PAL, DT y la percepción subjetiva del paso del tiempo). Consistencia interna: alfa de Cronbach. Estructura factorial analizada a través de un análisis factorial exploratorio. Precisión diagnóstica: análisis de área bajo la curva ROC, sensibilidad y especificidad.</p>	<p><b>Criterios prueba diagnóstica positiva:</b> Este estudio explora el punto de corte óptimo del ENP-E, por lo que en un principio no está predeterminado ni descrito. No se detectaron diferencias significativas en la presencia de preocupaciones entre pacientes con o sin necesidades psicológicas, por lo que la sección B no se tuvo en cuenta para la puntuación total. Sí existen diferencias significativas entre quienes muestran signos externos de malestar emocional (sección C), pero no se aclara cómo se contempla esto a la hora de establecer el punto de corte.</p> <p><b>Criterios prueba de comparación positiva:</b> La prueba utilizada como comparación fue una combinación de los valores de la HADS (HADS mayor o igual a 17), junto con la puntuación del ítem de calidad de vida del EORTC QLQ-C15-PAL (menor o igual a 3), la puntuación del DT (mayor a 5), y el grado de sufrimiento (percepción a largo plazo). Esta combinación se construyó para indicar si el valor en su conjunto es positivo o no. No queda claro cómo puntúa el grado de sufrimiento y a partir de qué puntuación se considera positivo.</p>			

## Anexo C. Referencias excluidas

CITA	CAUSA DE EXCLUSIÓN
Almonacid Guinot V. Psicoterapia en cáncer de mama: propuesta y resultados de un programa estructurado [Tesis Doctoral]. Valencia: Universidad de Valencia; 2014.	- Tesis cuyo objetivo no es la validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Andreu Y, Galdon MJ, Dura E, Martínez P, Perez S, Murgui S. A longitudinal study of psychosocial distress in breast cancer: Prevalence and risk factors. <i>Psychol Health</i> . 2012;27(1):72-87. PubMed PMID: 21678180.	- No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Belar A, Arantzamendi M, Rodríguez-Núñez A, Santesteban Y, Martínez M, López-Saca M, et al. Multicenter Study of the Psychometric Properties of the New Demoralization Scale (DS-II) in Spanish-Speaking Advanced Cancer Patients. <i>J Pain Symptom Manage</i> . 2019;57(3):627-34.	- Estudio de validación de una herramienta que evalúa otro constructo.
Bellver-Pérez A, Peris-Juan C, Santaballa-Beltrán A. Effectiveness of group therapy in women with localized breast cancer. <i>Int J Clin Health Psychol</i> . 2019;19(2):107-14	- No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Calderon C, Carmona-Bayonas A, Beato C, Ghanem I, Hernandez R, Majem M, et al. Risk of malnutrition and emotional distress as factors affecting health-related quality of life in patients with resected cancer. <i>Clin Transl Oncol</i> . 2019;21(5):687-91.	- No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Calderon C, Ferrando PJ, Lorenzo-Seva U, Higuera O, Ramon YCT, Rogado J, et al. Validity and Reliability of the Decision Regret Scale in Cancer Patients Receiving Adjuvant Chemotherapy. <i>J Pain Symptom Manage</i> . 2019;57(4):828-34.	- Estudio de validación de una herramienta que evalúa otro constructo.
Calderon C, Jimenez-Fonseca P, Hernandez R, Mar Muñoz MD, Mut M, Mangas-Izquierdo M, et al. Quality of life, coping, and psychological and physical symptoms after surgery for non-metastatic digestive tract cancer. <i>Surg Oncol</i> . 2019;31:26-32.	- No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Castellano Álvarez MR. El psicólogo en cuidados paliativos. Análisis de la demanda y de la eficacia de la intervención en la Región de Murcia [Tesis Doctoral]. Murcia: Universidad de Murcia; 2016.	- Tesis cuyo objetivo no es la validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Charlo Bernardos M. Crecimiento postraumático y factores psicosociales en mujeres con cáncer de mama vinculadas a asociación de pacientes [Tesis Doctoral]. Zaragoza: Universidad de Zaragoza; 2015.	- Tesis cuyo objetivo no es la validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Cruzado JA, Martínez García V, Salas Gutiérrez V, Jarabo Sarceda JR, Fraile Olivero CA, Fernández Martín E, et al. Implementación de un programa de cribado del malestar emocional en un servicio de cirugía torácica. <i>Cir Esp</i> . 2019;97(5):275-81.	- Estudio realizado con población oncológica y no oncológica.
de Linares Fernández S. Valoración de la influencia de variables psicológicas y estados emocionales en la implantación de progenitores hematopoyéticos en trasplante de médula. Efectos de un programa de intervención psicológica [Tesis Doctoral]. Granada: Universidad de Granada; 2014.	- Tesis cuyo objetivo no es la validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Díaz-Frutos D, Baca-García E, García-Foncillas J, LópezCastroman J. Predictors of psychological distress in advanced cancer patients under palliative treatments. <i>Eur J Cancer Care (Engl)</i> . 2016;25(4):608-15.	- No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Diego Pedro R. Acompanyament espiritual, malestar emocional i resiliència en pacients en situació de cures pal·liatives i efectes d'una intervenció en espiritualitat [Tesis Doctoral]. Valencia: Universitat de València; 2017.	- Tesis cuyo objetivo no es la validación de herramientas de cribado de malestar emocional en pacientes oncológicos. Lengua excluida.
Estepa T, Estepa J, Soria-Pastor S, Díez A. Uso de internet para evaluar el estrés psicológico en pacientes con cáncer de mama. <i>Psicooncología</i> . 2014;11(2-3):271-83.	- No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Fernández B, Bejar E, Campos MM. Primer impacto: programa de detección de estrés y atención psicosocial para paciente recién diagnosticado y sus familiares. <i>Psicooncología</i> . 2012;9(2-3):317-34.	- No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Fernández de Larrea-Baz N, Pérez-Gómez B, GuerreroZotano A, Casas AM, Bermejo B, Baena-Cañada JM, et al. Primary breast cancer and health related quality of life in Spanish women: The EpiGEICAM case-control study. <i>Sci Rep</i> . 2020;10(1).	- No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Gil F, Costa G, Hilker I, Benito L. First Anxiety, Afterwards Depression: Psychological Distress in Cancer Patients at Diagnosis and after Medical Treatment. <i>Stress Health</i> . 2012;28(5):362-7.	- No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.

CITA	CAUSA DE EXCLUSIÓN
Gil F, Grassi L, Travado L, Tomamichel M, Gonzalez JR, Southern European Psycho-Oncology Study G. Use of distress and depression thermometers to measure psychosocial morbidity among southern European cancer patients. <i>Support Care Cancer</i> . 2005;13(8):600-6. PubMed PMID: 15761700.	– Estudio realizado con pacientes oncológicos de Italia, Portugal, España y Suiza. No se ofrecen datos desagregados por país.
Gil Moncayo FL, Maté Méndez J, Ochoa Arnedo C, Casellas-Grau A, Trelis Navarro J, Borrás Andrés JM. Abordaje de la complejidad psicosocial en pacientes con cáncer. <i>Med Clin</i> . 2020;154(3):86-93.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Gomez D, Calderon C, Carmona-Bayonas A, Cacho Lavin D, Munoz MM, Martinez Cabanez R, et al. Impact of adjuvant therapy toxicity on quality of life and emotional symptoms in patients with colon cancer: a latent class analysis. <i>Clin Transl Oncol</i> . 2021;23(3):657-62. PubMed PMID: 32676895.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Gonzalez-Fernandez S, Fernandez-Rodriguez C, Mota-Alonso M, Garcia-Tejido P, Pedrosa I, Perez-Alvarez M. Emotional state and psychological flexibility in breast cancer survivors. <i>Eur J Oncol Nurs</i> . 2017;30:75-83.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Gonzalez-Saenz de Tejada M, Bilbao A, Baré M, Briones E, Sarasqueta C, Quintana JM, et al. Association of social support, functional status, and psychological variables with changes in health-related quality of life outcomes in patients with colorectal cancer. <i>Psycho-oncology</i> . 2016;25(8):891-7.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Grassi L, Berardi MA, Ruffilli F, Meggiolaro E, Andritsch E, Sirgo A, et al. Role of psychosocial variables on chemotherapy-induced nausea and vomiting and health-related quality of life among cancer patients: a European study. <i>Psychother Psychosom</i> . 2015;84(6):339-47.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Jimenez-Fonseca P, Calderon C, Carmona-Bayonas A, Munoz MM, Hernandez R, Mut Lloret M, et al. The relationship between physician and cancer patient when initiating adjuvant treatment and its association with sociodemographic and clinical variables. <i>Clin Transl Oncol</i> . 2018;20(11):1392-9.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Jimenez-Fonseca P, Lorenzo-Seva U, Ferrando PJ, Carmona-Bayonas A, Beato C, Garcia T, et al. The mediating role of spirituality (meaning, peace, faith) between psychological distress and mental adjustment in cancer patients. <i>Support Care Cancer</i> . 2018;26(5):1411-8.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Martinez Arroyo O, Andreu Vaillo Y, Martinez Lopez P, Galdon Garrido MJ. Emotional distress and unmet supportive care needs in survivors of breast cancer beyond the end of primary treatment. <i>Support Care Cancer</i> . 2019;27(3):1049-57.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Martínez López P. Cribado de distrés emocional en población oncológica adulta: validación de instrumentos [Tesis Doctoral]. Valencia: Universidad de Valencia; 2016.	– Tesis por compendio de artículos. Se seleccionaron aquellos artículos de validación que todavía no habían sido localizados y que cumplen los criterios de inclusión/exclusión (3).
Martinez P, Andreu Y, Conchado A. Psychometric Properties of the Spanish Version of the Cancer Survivors' Unmet Needs (CaSUN-S) Measure in Breast Cancer. <i>Psicothema</i> . 2021;33(1):155-63.	– Estudio de validación de una herramienta que evalúa otro constructo.
Maté Méndez J. Sufrimiento en el paciente oncológico al final de la vida [Tesis Doctoral]. Barcelona: Universidad Autónoma de Barcelona; 2014.	– Tesis por compendio de artículos. Se seleccionaron aquellos artículos de validación que todavía no habían sido localizados y que cumplen los criterios de inclusión/exclusión (1).
Montesinos F, Luciano C. Acceptance of relapse fears in breast cancer patients: effects of an ACT-based abridged intervention. <i>Psicooncología</i> . 2016;13(1):7-21.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Navarro Jiménez JM. La concordancia de la información sobre la enfermedad entre el paciente con cáncer avanzado y su cuidador principal [Tesis Doctoral]. Madrid: Universidad Complutense; 2019.	– Tesis cuyo objetivo no es la validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Ochoa C, Casellas-Grau A, Vives J, Font A, Borrás JM. Positive psychotherapy for distressed cancer survivors: Posttraumatic growth facilitation reduces posttraumatic stress. <i>Int J Clin Health Psychol</i> . 2017;17(1):28-37.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Ochoa-Arnedo C, Medina JC, Flix-Valle A, Anastasiadou D. E-health ecosystem with integrated and stepped psychosocial services for breast cancer survivors: study protocol of a multicentre randomised controlled trial. <i>BMJ open</i> . 2021;11(3):e041548. PubMed PMID: 34006024.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.

CITA	CAUSA DE EXCLUSIÓN
Perez-Tejada J, Garmendia L, Labaka A, Vegas O, GómezLazaro E, Arregi A. Active and Passive Coping Strategies: Comparing psychological distress, cortisol, and proinflammatory cytokine levels in breast cancer survivors. Clin J Oncol Nurs. 2019;23(6):583-90.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Petrova D, Redondo-Sanchez D, Rodriguez-Barranco M, Romero Ruiz A, Catena A, Garcia-Retamero R, et al. Physical comorbidities as a marker for high risk of psychological distress in cancer patients. Psychooncology. 2021;30(7):1160-1166. PubMed PMID: 33599019.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Prieto Díaz A. Factores cognitivos asociados con el desarrollo de síntomas ansioso-depresivos en pacientes oncológicos recién diagnosticados [Tesis Doctoral]. Santander: Universidad de Cantabria; 2016.	– Tesis cuyo objetivo no es la validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Rodríguez Torres J, Cabrera Martos I, López López L, Torres Sánchez I, Granados Santiago M, Carmen Valenza M. Psychological distress at hospital admission is related to symptoms severity and health status in malignant pleural effusion patients. Eur J Cancer Care (Engl). 2020;29(2). PubMed PMID: 31880035.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Romero R, Penas R, Garcia-Conde A, Llombart P, Perez S, Marco JH. Mindfulness-Based Cognitive Therapy for Spanish Oncology Patients: The Bartley Protocol. Psicothema. 2020;32(4):508-15.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Soto Rubio AL. Beneficios de un programa de soporte para familiares de pacientes al final de la vida: estudio multicéntrico [Tesis Doctoral]. Valencia: Universidad de Valencia; 2018.	– Tesis cuyo objetivo no es la validación de herramientas de cribado de malestar emocional en pacientes oncológicos.
Yélamos Agua C, Berzal Perez E, de Haro Gazquez D, Fernandez Sanchez B, Navarro Jimenez JM. Psychosocial impact of the COVID-19 outbreak and lockdown on Spanish oncological patients: a quantitative approach. J Psychosoc Oncol. 2021;39(3):385-98. PubMed PMID: 33570014.	– No es un estudio de validación de herramientas de cribado de malestar emocional en pacientes oncológicos.

## Anexo D. Evaluación de la calidad de los estudios de precisión diagnóstica: herramienta QUADAS-2

### D.1 Herramienta QUADAS-2

<b>DOMINIO 1: Selección de pacientes</b>	
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>	<input type="checkbox"/> Bajo <input type="checkbox"/> Alto <input type="checkbox"/> Incierto
<i>Describir el método de selección de pacientes:</i>	
<i>Preguntas orientativas:</i>	
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?	<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Incierto
– ¿se evitó un diseño de casos y controles?	<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Incierto
– ¿se evitaron exclusiones inapropiadas?	<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Incierto
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>	<input type="checkbox"/> Baja <input type="checkbox"/> Alta <input type="checkbox"/> Incierta
<b>DOMINIO 2: Prueba índice</b>	
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>	<input type="checkbox"/> Bajo <input type="checkbox"/> Alto <input type="checkbox"/> Incierto
<i>Describir cómo fue realizada e interpretada la prueba índice:</i>	
<i>Preguntas orientativas:</i>	
– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?	<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Incierto
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>	<input type="checkbox"/> Baja <input type="checkbox"/> Alta <input type="checkbox"/> Incierta
<b>DOMINIO 3: Prueba de referencia</b>	
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>	<input type="checkbox"/> Bajo <input type="checkbox"/> Alto <input type="checkbox"/> Incierto
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i>	
<i>Preguntas orientativas:</i>	
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?	<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Incierto
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?	<input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> Incierto

<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>	<input type="checkbox"/> Baja
	<input type="checkbox"/> Alta
	<input type="checkbox"/> Incierta
<b>DOMINIO 4: Flujo y tiempos</b>	
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>	<input type="checkbox"/> Bajo
	<input type="checkbox"/> Alto
	<input type="checkbox"/> Incierto
<i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2: Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i>	
<i>Preguntas orientativas:</i>	
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?	<input type="checkbox"/> Sí
	<input type="checkbox"/> No
	<input type="checkbox"/> Incierto
– ¿recibieron todos los pacientes la misma prueba de referencia?	<input type="checkbox"/> Sí
	<input type="checkbox"/> No
	<input type="checkbox"/> Incierto
– ¿se incorporaron todos los pacientes al análisis?	<input type="checkbox"/> Sí
	<input type="checkbox"/> No
	<input type="checkbox"/> Incierto
<b>Fuente:</b> Ciapponi A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.	

## D.2 Evaluación de la calidad de los estudios incluidos mediante la herramienta QUADAS-2

Herramienta	Estudio	Riesgo de sesgo				Preocupaciones sobre la aplicabilidad		
		Selección de pacientes	Prueba índice	Prueba de referencia	Flujo y tiempos	Selección de pacientes	Prueba índice	Prueba de referencia
BSI-18	Galdón, 2008	?	☺	–	☺	☺	☺	–
	Martínez, 2019	☺	☺	–	☺	☺	☺	–
	Calderon, 2020	☺	☺	–	☺	☺	☺	–
ESAS-r	Carvajal, 2013	?	☺	–	?	☹	☹	–
EVA	Martínez, 2006	☺	☺	☺	☺	☺	☺	☺
HADS	López-Roig, 2000	?	?	?	☹	☺	☺	?
	Costa, 2009	?	☺	?	☺	☺	☺	☺
	Martínez, 2012	☺	☺	☺	☺	☺	☺	☺
DT	Martínez, 2013	☺	☺	☺	☺	☺	☺	☺
	Priede, 2014	?	☺	☹	☹	?	☺	☺
	Martínez, 2015	☺	☺	☺	☺	☺	☺	☺
	Andreu, 2015	☺	☺	☺	☺	☺	☺	☺
	Andreu, 2019	☺	☺	☺	☺	☺	☺	☺
DME	Limonero, 2012	☺	☹	☺	☺	☺	☺	☺
ENP-E	Mateo-Ortega, 2019	?	☺	?	☺	☺	☺	☺

Valoración de las preocupaciones respecto al riesgo de sesgo y a la aplicabilidad: ☺ Riesgo bajo; ☹ Riesgo alto; ? Riesgo incierto.

**Abreviaturas:** BSI-18 = Inventario Breve de Síntomas18; ESAS-r = Cuestionario de evaluación de síntomas de Edmonton, versión revisada; EVA de an y de = Escalas Visuales Analógicas de ansiedad y depresión; HADS = Escala Hospitalaria de Ansiedad y Depresión; DT = Termómetro de distrés; DME = Cuestionario de Detección de Malestar Emocional; ENP-E = Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida.

## Anexo E. Valoración individual del riesgo de sesgo y la aplicabilidad de cada estudio mediante la herramienta QUADAS-2

Autoría/año: Galdón, 2008 (90)	Herramienta: BSI-18	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Incierto
<i>Describir el método de selección de pacientes:</i>		
La muestra se obtuvo de 295 mujeres con resultados anormales en una mamografía, que habían sido derivadas a la Fundación Instituto Valenciano de Oncología (Valencia, España) para establecer su diagnóstico y, de ser necesario, recibir tratamiento. Las participantes elegibles debían ser mujeres que cumplan los siguientes requisitos: (a) tener 18 años o más; (b) haber recibido un diagnóstico primario de cáncer de mama; (c) no estar participando en ningún ensayo clínico; (d) proporcionar el consentimiento informado. De esta muestra inicial, 120 mujeres no cumplieron los criterios de inclusión (89 acabaron no teniendo cáncer, 17 estaban participando en un ensayo clínico y 14 rechazaron formar parte del estudio). La muestra final estuvo formada por 175 mujeres que habían sido diagnosticadas recientemente de cáncer de mama. El rango de edad fue de 27 a 70 años, y la media de edad 52,7 años (dt = 10,12).		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Incierto
– ¿se evitó un diseño de casos y controles?		Sí
– ¿se evitaron exclusiones inapropiadas?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cómo fue realizada e interpretada la prueba índice:</i>		
La prueba índice es el BSI-18. Una vez se confirmó el diagnóstico, y antes de comenzar el tratamiento posquirúrgico, las participantes completaron un cuestionario autoadministrado que incluía datos demográficos y la versión española del BSI-18. La interpretación de la prueba índice puede considerarse objetiva, dado que fue autoadministrada y las puntuaciones de cada ítem se aplican según la respuesta directa aportada por la paciente. Posteriormente, se pasan a una escala T estandarizada para interpretarlas. El punto de corte de la escala se explicita con anterioridad.		
<i>Preguntas orientativas:</i>		
– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		-
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i>		
No se realiza prueba de referencia.		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?		-
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?		-
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		-
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i>		
De la muestra inicial (295 mujeres), 120 mujeres no cumplieron los criterios de inclusión (89 acabaron no teniendo cáncer, 17 estaban participando en un ensayo clínico, y 14 rechazaron formar parte del estudio). Se desconoce cuáles son las características de las pacientes que no han llegado a participar en el estudio.		
<i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i>		
Al no realizarse la prueba de referencia, no existe intervalo de tiempo entre la prueba índice y la de referencia.		

Autoría/año: Galdón, 2008 (90)	Herramienta: BSI-18	Valoración
<i>Preguntas orientativas:</i>		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?		-
– ¿recibieron todos los pacientes la misma prueba de referencia?		-
– ¿se incorporaron todos los pacientes al análisis?		Sí
<b>Fuente:</b> Ciapponi A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.		

Autoría/año: Martínez, 2019 (92)	Herramienta: BSI-18	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<i>Describir el método de selección de pacientes:</i>		
La población objeto de estudio son pacientes oncológicos en edad adulta. Los y las participantes elegibles son pacientes que cumplan los siguientes requisitos: (a) tener 18 años o más; (b) tener un diagnóstico de cáncer; (c) poder proporcionar el consentimiento informado. Se invitó a participar a las y los pacientes consecutivos que visitaron la Fundación Instituto Valenciano de Oncología (FIVO) para recibir atención en cualquier departamento médico entre mayo de 2015 y julio de 2017. Del total de 1259 pacientes que se abordaron, 76 (6 %) declinaron participar. La muestra final estuvo formada por 1183 pacientes oncológicos (65 % mujeres). La media de edad de los y las participantes fue 57,71 años (dt = 13,09; rango de edad = 19-94 años).		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Sí
– ¿se evitó un diseño de casos y controles?		Sí
– ¿se evitaron exclusiones inapropiadas?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> <b>Baja</b>
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<i>Describir cómo fue realizada e interpretada la prueba índice:</i>		
La prueba índice es el BSI-18. Los y las pacientes recibieron información sobre el estudio y, quienes tenían interés en participar, proporcionaron el consentimiento informado y completaron el conjunto de cuestionarios de manera autoadministrada. La interpretación de la prueba índice puede considerarse objetiva, dado que fue autoadministrada y las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente. Posteriormente, se pasan a una escala T estandarizada. El punto de corte de la escala se explicita con anterioridad.		
<i>Preguntas orientativas:</i>		
– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> <b>Baja</b>
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		-
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i>		
No se realiza prueba de referencia.		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?		-
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?		-
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		-

Autoría/año: Martínez, 2019 (92)	Herramienta: BSI-18	Valoración
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<p>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2: Del total de 1259 pacientes que se abordaron, 76 (6 %) declinaron participar. Se desconoce cuáles son las características de las y los pacientes que no han llegado a participar en el estudio.</p> <p><i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i> Al no realizarse la prueba de referencia, no existe intervalo de tiempo entre la prueba índice y la de referencia.</p> <p><i>Preguntas orientativas:</i></p> <ul style="list-style-type: none"> <li>- ¿hubo un intervalo apropiado entre la prueba índice y la de referencia? -</li> <li>- ¿recibieron todos los pacientes la misma prueba de referencia? -</li> <li>- ¿se incorporaron todos los pacientes al análisis? Sí</li> </ul> <p><b>Fuente:</b> Ciapponi A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.</p>		

Autoría/año: Calderon, 2020 (93)	Herramienta: BSI-18	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<p><i>Describir el método de selección de pacientes:</i> La población consiste en pacientes con cáncer. La muestra fue consecutiva y se seleccionó a las y los pacientes con cáncer que acudieron a alguno de los 10 centros participantes en el estudio entre julio de 2015 y febrero de 2019. Los pacientes tenían un diagnóstico de cáncer de mama o cáncer colorrectal confirmado histológicamente, no avanzado y tratado quirúrgicamente, para los que las directrices clínicas internacionales consideraban que el tratamiento adyuvante era una opción. Se excluyó a los y las pacientes que eran menores de 18 años, habían recibido tratamiento con radio o quimioterapia preoperatoria, o solo con terapia hormonal o con radioterapia adyuvante sin quimioterapia, o tenían alguna enfermedad mental grave que les impidiera comprender el estudio. Se seleccionaron 1003 pacientes, de entre quienes se excluyó a 126. La muestra final estuvo formada por 877 pacientes, con una media de edad de 36 años (dt = 13,4; rango entre 25 y 84 años), y el 59,7 % eran mujeres.</p> <p><i>Preguntas orientativas:</i></p> <ul style="list-style-type: none"> <li>- ¿se incluyó una muestra consecutiva o aleatoria de pacientes? Sí</li> <li>- ¿se evitó un diseño de casos y controles? Sí</li> <li>- ¿se evitaron exclusiones inapropiadas? Sí</li> </ul> <p><b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b></p>		
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<p><i>Describir cómo fue realizada e interpretada la prueba índice:</i> La prueba índice es el BSI-18. Los y las pacientes completaron este instrumento de manera autoadministrada durante la semana siguiente a su primera visita al departamento de oncología médica para decidir el tratamiento adyuvante. La interpretación de la prueba índice puede considerarse objetiva, dado que fue autoadministrada y las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente. Posteriormente, se pasan a una escala T estandarizada. El punto de corte de la escala se explicita con anterioridad.</p> <p><i>Preguntas orientativas:</i></p> <ul style="list-style-type: none"> <li>- ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia? Sí</li> <li>- <b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b></li> </ul>		
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		-

Autoría/año: Calderon, 2020 (93)	Herramienta: BSI-18	Valoración
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i> No se realiza prueba de referencia.		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?		-
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?		-
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		-
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i> Del total de 1003 pacientes que se seleccionaron, se excluyeron 126. Se desconoce cuáles son las características de las y los pacientes que no han llegado a participar en el estudio. <i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i> Al no realizarse la prueba de referencia, no existe intervalo de tiempo entre la prueba índice y la de referencia.		
<i>Preguntas orientativas:</i>		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?		-
– ¿recibieron todos los pacientes la misma prueba de referencia?		-
– ¿se incorporaron todos los pacientes al análisis?		Sí
<b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.		

Autoría/año: Carvajal, 2013 (100)	Herramienta: ESAS-r	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Incierto
<i>Describir el método de selección de pacientes:</i> La población objeto de estudio son pacientes con cáncer avanzado. Todas y todos los pacientes eran mayores de 18 años, con una función cognitiva normal en el Mini-Mental State Examination (MMSE). Los y las pacientes con cáncer avanzado que estaban recibiendo radioterapia paliativa o quimioterapia también eran elegibles. La muestra total de pacientes para el estudio psicométrico del ESAS-r fue de 66 pacientes (al menos 30 de cada centro participante). El 70 % eran mujeres, y la edad media fue de 54 años (con un rango de 18 a 84 años). No se menciona si la muestra es consecutiva o aleatoria.		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Incierto
– ¿se evitó un diseño de casos y controles?		Sí
– ¿se evitaron exclusiones inapropiadas?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Alta
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cómo fue realizada e interpretada la prueba índice:</i> La prueba índice es el ESAS-r y fue autoinformada por parte de los y las pacientes, en presencia del personal investigador. La interpretación de la prueba índice puede considerarse objetiva, dado que las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente. A lo largo del artículo no se explicita el punto de corte de cada uno de los ítems o el total de la escala.		
<i>Preguntas orientativas:</i>		
– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Alta

Autoría/año: Carvajal, 2013 (100)	Herramienta: ESAS-r	Valoración
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		-
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i> No se realiza prueba de referencia.		
<i>Preguntas orientativas:</i>		
- ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?		-
- ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?		-
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		-
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Incierto
<i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i> Todas y todos los pacientes de la muestra recibieron la prueba índice, pero nadie recibió una prueba de referencia. Se desconoce si hubo pacientes que declinaron participar en el estudio y, por tanto, se desconocen sus características.		
<i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i> Al no realizarse la prueba de referencia, no existe intervalo de tiempo entre la prueba índice y la de referencia.		
<i>Preguntas orientativas:</i>		
- ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?		-
- ¿recibieron todos los pacientes la misma prueba de referencia?		-
- ¿se incorporaron todos los pacientes al análisis?		Incierto
<b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.		

Autoría/año: Martínez, 2006 (101)	Herramienta: EVA de ansiedad y depresión	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir el método de selección de pacientes:</i> La población objeto de estudio son pacientes con diagnóstico de cáncer avanzado o terminal a quienes atienden los equipos de cuidados paliativos. Como criterios de inclusión se determinaron: (1) pacientes con diagnóstico de cáncer; (2) mayores de 18 años; (3) pacientes que acuden por primera vez a la consulta externa (CEXT) del Servicio de Cuidados Paliativos (SCP) del Instituto Catalán de Oncología (ICO). Como criterio de exclusión, se estableció la presencia de un deterioro del estado general del paciente que, a criterio profesional, no permitiera la evaluación y la valoración completa de los diferentes parámetros que se medían en el estudio. Se trata de una muestra consecutiva de pacientes que acuden por primera vez a la consulta externa del Servicio de Cuidados Paliativos del Instituto Catalán de Oncología desde septiembre a noviembre de 2001. La muestra inicial fue de 54 pacientes. Un paciente no pudo ser evaluado debido a su discapacidad intelectual. La muestra final de pacientes fue de 53, el 66 % hombres. Edad media 63 años (dt = 12,59).		
<i>Preguntas orientativas:</i>		
- ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Sí
- ¿se evitó un diseño de casos y controles?		Sí
- ¿se evitaron exclusiones inapropiadas?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo

Autoría/año: Martínez, 2006 (101)	Herramienta: EVA de ansiedad y depresión	Valoración
<i>Describir cómo fue realizada e interpretada la prueba índice:</i>		
Las pruebas índice son las Escalas Visuales Analógicas (EVA) de ansiedad y depresión. Estas pruebas fueron administradas por parte del personal médico o de enfermería en la primera visita a la CEXT del SCP del ICO. El personal médico daba instrucciones a los y las pacientes sobre la ejecución de las EVA. No se aclara el rango de las EVA, pero se entiende que sería de 0 a 10. El sumatorio de las puntuaciones de ambas EVA da como resultado el Distrés Emocional Global (DEG). La interpretación de la prueba índice puede considerarse objetiva, dado que la puntuación de cada escala es la respuesta directa señalada por el o la paciente.		
<i>Preguntas orientativas:</i>		
– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?	Sí	
– <b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>	<input checked="" type="checkbox"/> Baja	
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>	-	
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i>		
La prueba de referencia es la escala HADS. En el mismo día y después de la visita médica, los y las pacientes tuvieron una primera entrevista con la psicóloga y se llevó a cabo una valoración psicológica que incluía la administración de la HADS. El punto de corte de la escala fue 16. La interpretación de la prueba de referencia puede considerarse objetiva, ya que la puntuación de cada ítem es la respuesta directa aportada por el o la paciente, y el personal que administra la prueba índice y la prueba de referencia es diferente.		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?	-	
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?	-	
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>	-	
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>	<input checked="" type="checkbox"/> Bajo	
<i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i>		
Solo se excluyó a un paciente (el 1,9 % de la muestra inicial). Se desconoce cuáles son sus características.		
<i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i>		
Primero, el personal médico y de enfermería administró la prueba índice. En el mismo día y posteriormente, la psicóloga del servicio administró la prueba de referencia. No existe ninguna intervención en este ámbito entre ambas aplicaciones.		
<i>Preguntas orientativas:</i>		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?	-	
– ¿recibieron todos los pacientes la misma prueba de referencia?	-	
– ¿se incorporaron todos los pacientes al análisis?	Sí	
<b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.		

Autoría/año: López-Roig, 2000 (94)	Herramienta: HADS	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>	<input checked="" type="checkbox"/> Incierto	
<i>Describir el método de selección de pacientes:</i>		
La población objeto de estudio es la población española de pacientes con cáncer. La muestra estudiada incluye 484 pacientes con cáncer, 336 mujeres (69 %) y 148 hombres (31 %). Los y las pacientes se entrevistaron en 7 hospitales o centros participantes en el territorio español, cuando acudían a las consultas externas: 139 pacientes (28,7 %); al hospital de día: 131 (27,1 %); a la planta hospitalaria: 112 (23,1 %); o a consulta de psicología: 102 (21,1 %). No se menciona si la muestra es consecutiva o aleatoria. Tampoco se explicita si hubo alguna exclusión en la selección de pacientes. No obstante, el hecho de reclutar pacientes en las consultas de psicología (21 % de la muestra) pudiera dar lugar a una sobreestimación de la precisión diagnóstica.		
<i>Preguntas orientativas:</i>		

Autoría/año: López-Roig, 2000 (94)	Herramienta: HADS	Valoración
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Incierto
– ¿se evitó un diseño de casos y controles?		Sí
– ¿se evitaron exclusiones inapropiadas?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Incierto</b>
<p><i>Describir cómo fue realizada e interpretada la prueba índice:</i></p> <p>No se indica si la prueba índice se llevó a cabo antes o después de la prueba de referencia. Se informa de que se adaptó el diseño para permitir su administración tanto en formato autoaplicado como de entrevista, pero se desconoce a cuántas personas se les aplicó de una u otra manera. Se intuye que ambas pruebas se realizaron el mismo día, pero tampoco se explicita como tal. No obstante, la interpretación de la prueba índice puede considerarse objetiva, dado que las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente. El equipo autor apoya el uso de las dos subescalas por separado, como medidas diferentes de ansiedad y depresión, aunque puedan coexistir en el tiempo, y no como una medida global de estrés.</p>		
<i>Preguntas orientativas:</i>		
¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?		Incierto
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> <b>Baja</b>
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Incierto</b>
<p><i>Describir cómo fue realizada e interpretada la prueba de referencia:</i></p> <p>No se indica si las pruebas de referencia se realizaron antes o después de la prueba índice. Se desconoce si se aplicaron en formato autoadministrado o heteroadministrado. Se intuye que tanto la prueba índice como las pruebas de referencia se efectuaron el mismo día, pero tampoco se explicita como tal. La interpretación de las pruebas de referencia puede considerarse objetiva, dado que las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente. Los puntos de corte para cada una de las escalas se explicitan con anterioridad. No se aportan datos de estudios previos que avalen la fiabilidad, la validez y la precisión diagnóstica de las pruebas de referencia (STAI-S; SDS).</p>		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?		Incierto
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?		Incierto
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		<input checked="" type="checkbox"/> <b>Incierta</b>
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Alto</b>
<p><i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i></p> <p>Todas y todos los pacientes de la muestra recibieron la prueba índice. Se administraron pruebas de referencia a 140 y 87 pacientes, respectivamente. Se desconoce cuáles son las características de los y las pacientes que no han recibido las pruebas de referencia.</p>		
<p><i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i></p> <p>No se explicita qué intervalo de tiempo pasa entre la aplicación de la prueba índice y las pruebas de referencia. Se desconoce si hubo algún tipo de intervención entre la prueba índice y las de referencia, pero se intuye que fueron administradas de manera continua.</p>		
<i>Preguntas orientativas:</i>		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?		Incierto
– ¿recibieron todos los pacientes la misma prueba de referencia?		No
– ¿se incorporaron todos los pacientes al análisis?		No
<p><b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.</p>		

Autoría/año: Costa, 2009 (95)	Herramienta: HADS	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Incierto</b>
<i>Describir el método de selección de pacientes:</i>		
La población son pacientes adultos oncológicos ambulatorios. Los criterios de inclusión fueron: a) ser paciente ambulatorio con diagnóstico de enfermedad oncológica y b) tener una edad comprendida entre los 18 y los 80 años. Los criterios de exclusión fueron: a) presentar trastorno psicótico y b) tener una capacidad funcional menor del 40 % (según el Karnofsky Performance Status). También se consideró como criterio de exclusión: c) presentar deterioro cognitivo que pudiera interferir en la comprensión de las pruebas administradas. Se solicitó la participación de 245 pacientes oncológicos ambulatorios, que acudían al hospital de día para recibir tratamiento de quimioterapia o asistían a las consultas en las unidades funcionales. El 95,9 % aceptaron participar, lo que significa que la muestra final estuvo compuesta por 234 pacientes. Del total, 121 eran mujeres (51,7 %). No se explicita si la muestra es consecutiva o aleatoria.		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?	Incierto	
– ¿se evitó un diseño de casos y controles?	Sí	
– ¿se evitaron exclusiones inapropiadas?	Incierto	
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> <b>Baja</b>
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<i>Describir cómo fue realizada e interpretada la prueba índice:</i>		
Los y las pacientes participantes fueron entrevistados para recoger datos sociodemográficos y médicos. Tras esa recogida de datos, una psicóloga clínica realizó la entrevista psiquiátrica SCID-I (prueba de referencia) y, después, los y las pacientes cumplimentaron la HADS (prueba índice). La interpretación de la prueba índice puede considerarse objetiva, dado que fue autoadministrada, y las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente.		
<i>Preguntas orientativas:</i>		
– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?	Incierto	
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Incierto</b>
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i>		
Las y los pacientes participantes fueron entrevistados para recoger datos sociodemográficos y médicos. Tras esa recogida de datos, una psicóloga clínica realizó la entrevista psiquiátrica SCID-I (prueba de referencia). Se administraron las preguntas de cribado de cada módulo conforme a lo que se propone en el cuaderno de aplicación del SCID-I (First et al., 1999). La entrevista se llevó a cabo en una única sesión con una duración aproximada de 45 a 90 minutos. Los diagnósticos obtenidos con la entrevista psiquiátrica se agruparon, según la sintomatología, en cuatro categorías: a) sintomatología depresiva; b) sintomatología de ansiedad; c) sintomatología mixta (ansiedad-depresión); d) pacientes sin trastorno psiquiátrico. Se desconoce si los resultados de la prueba de referencia se interpretaron con o sin conocimiento previo de los resultados de la prueba índice.		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?	Sí	
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?	Incierto	
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		<input checked="" type="checkbox"/> <b>Baja</b>
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i>		
Todas y todos los pacientes participantes de la muestra recibieron la prueba índice y la de referencia.		
<i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i>		
Se aplica la prueba de referencia de manera heteroadministrada y, acto seguido, se aplica la prueba índice de manera autoadministrada.		

Autoría/año: Costa, 2009 (95)	Herramienta: HADS	Valoración
<i>Preguntas orientativas:</i>		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?		Sí
– ¿recibieron todos los pacientes la misma prueba de referencia?		Sí
– ¿se incorporaron todos los pacientes al análisis?		Sí
<b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.		

Autoría/año: Martínez, 2012 (96)	Herramienta: HADS	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<i>Describir el método de selección de pacientes:</i>		
La población objeto de estudio son pacientes adultos oncológicos. Los criterios de inclusión fueron: a) tener 18 años o más; b) tener un diagnóstico de cáncer; y c) ser capaz de proporcionar el consentimiento informado. Se trata de una muestra consecutiva de pacientes que visitaron la Fundación Instituto Valenciano de Oncología (FIVO) desde octubre de 2010 a abril de 2012. Del total de 945 pacientes que se abordaron, el 6 % declinó participar en el estudio. La muestra final estuvo formada por 892 pacientes oncológicos, el 70 % mujeres.		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Sí
– ¿se evitó un diseño de casos y controles?		Sí
– ¿se evitaron exclusiones inapropiadas?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> <b>Baja</b>
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<i>Describir cómo fue realizada e interpretada la prueba índice:</i>		
Después de recibir información sobre el estudio y proporcionar el consentimiento informado, las y los participantes completaron un cuestionario autoadministrado que incluía un apartado de datos demográficos, la prueba índice (HADS) y la prueba de referencia (BSI-18). La interpretación de la prueba índice puede considerarse objetiva, dado que fue autoadministrada y las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente.		
<i>Preguntas orientativas:</i>		
¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> <b>Baja</b>
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i>		
Después de recibir información sobre el estudio y proporcionar el consentimiento informado, los y las participantes completaron un cuestionario autoadministrado que incluía un apartado de datos demográficos, la prueba índice (HADS) y la prueba de referencia (BSI-18). La interpretación de la prueba de referencia puede considerarse objetiva, dado que fue autoadministrada y las puntuaciones de cada ítem se aplica según la respuesta directa aportada por el o la paciente. Posteriormente, se pasaron a una escala T estandarizada. El punto de corte de la escala se explicita con anterioridad. Se aportan datos de estudios previos que avalan la fiabilidad y la validez de la prueba de referencia.		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?		Sí
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?		Si
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		<input checked="" type="checkbox"/> <b>Baja</b>

Autoría/año: Martínez, 2012 (96)	Herramienta: HADS	Valoración
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i> Hubo 53 personas que declinaron participar en el estudio (el 6 % de la muestra inicial). Las razones esgrimidas fueron: no tener interés (45,3 %), sentir demasiado cansancio (28,3 %), tener prisa (9,4 %) y otras razones (17 %).		
<i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i> Tanto la prueba índice como la prueba de referencia se administraron consecutivamente, pues forman parte del cuestionario autoadministrado que ha de cubrir cada participante.		
<i>Preguntas orientativas:</i>		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?		Sí
– ¿recibieron todos los pacientes la misma prueba de referencia?		Sí
– ¿se incorporaron todos los pacientes al análisis?		Sí
<b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.		

Autoría/año: Martínez, 2013 (97)	Herramienta: DT	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir el método de selección de pacientes:</i> La población objeto de estudio son pacientes adultos oncológicos. Los criterios de inclusión fueron: a) tener más de 18 años; b) tener un diagnóstico de cáncer; y c) ser capaz de proporcionar el consentimiento informado. Se trata de una muestra consecutiva de pacientes que visitaron la Fundación Instituto Valenciano de Oncología (FIVO) para recibir tratamiento en alguno de sus departamentos, desde octubre de 2010 a octubre de 2012. Del total de 1023 pacientes que se abordaron, el 6 % declinó participar en el estudio. La muestra final estuvo formada por 962 participantes, el 68 % mujeres.		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Sí
– ¿se evitó un diseño de casos y controles?		Sí
– ¿se evitaron exclusiones inapropiadas?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Bajo
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cómo fue realizada e interpretada la prueba índice:</i> Después de recibir información sobre el estudio y proporcionar el consentimiento informado, los y las participantes completaron un cuestionario autoadministrado que incluía un apartado de datos, la prueba índice (DT) la Lista de Problemas (PL), la prueba de referencia (BSI-18) y un cuestionario psicosocial sobre el conocimiento y el uso del departamento de psicooncología por parte de los y las pacientes. La interpretación de la prueba índice puede considerarse objetiva, dado que fue autoadministrada y la puntuación señalada por el o la paciente se interpretó directamente.		
<i>Preguntas orientativas:</i>		
¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo

Autoría/año: Martínez, 2013 (97)	Herramienta: DT	Valoración
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i>		
Después de recibir información sobre el estudio y proporcionar el consentimiento informado, los y las participantes completaron un cuestionario autoadministrado que incluía un apartado de datos, la prueba índice (DT), la Lista de Problemas (PL), la prueba de referencia (BSI-18) y un cuestionario psicosocial sobre el conocimiento y el uso del departamento de psicooncología por parte de los y las pacientes. La interpretación de la prueba de referencia puede considerarse objetiva, dado que fue autoadministrada y las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente. Posteriormente, se pasan a una escala T estandarizada. El punto de corte de la escala se explicita con anterioridad. Se aportan datos de estudios previos que avalan las propiedades psicométricas adecuadas de la prueba de referencia.		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?		Sí
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?		Si
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i>		
Hubo 61 personas que declinaron participar en el estudio (el 6 % de la muestra inicial). Las razones esgrimidas fueron: no tener interés (45 %), sentir demasiado cansancio (28 %), tener prisa (10 %) y otras razones (17 %).		
<i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i>		
Tanto la prueba índice como la prueba de referencia se administraron consecutivamente, pues forman parte del cuestionario autoadministrado que ha de cubrir cada participante.		
<i>Preguntas orientativas:</i>		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?		Sí
– ¿recibieron todos los pacientes la misma prueba de referencia?		Sí
– ¿se incorporaron todos los pacientes al análisis?		Sí
<b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.		

Autoría/año: Priede, 2014 (91)	Herramienta: DT	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Incierto
<i>Describir el método de selección de pacientes:</i>		
La población objeto de estudio son pacientes en edad adulta con un primer diagnóstico reciente de cáncer. El personal facultativo de los servicios implicados en la atención médica de pacientes con cáncer del Hospital Universitario Marqués de Valdecilla (Santander, España) derivó a 276 pacientes al programa de investigación donde se enmarca este estudio. Estas personas contaban con un primer diagnóstico reciente de cáncer confirmado por estudio de anatomía patológica. Se excluyó a 59 pacientes, y los criterios de exclusión fueron: a) pacientes con un tiempo desde el diagnóstico mayor a 4 meses; b) pacientes menores de 18 años; c) pacientes mayores de 70 años; d) pacientes con recidivas o que hayan padecido otro proceso oncológico con anterioridad; e) pacientes con discapacidad intelectual, deterioro de las funciones cognitivas o incapacidad para comprender las pruebas de evaluación. Del conjunto de pacientes que se incluyeron (217), 22 rechazaron participar. De 195 pacientes que accedieron a participar en el estudio, 135 realizaron la evaluación y entregaron los cuestionarios cumplimentados. De ese total, 88 eran mujeres (65,19 %). No se explicita si la muestra es consecutiva o aleatoria.		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Incierto
– ¿se evitó un diseño de casos y controles?		Sí
– ¿se evitaron exclusiones inapropiadas?		Incierto
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Incierto
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo

Autoría/año: Priede, 2014 (91)	Herramienta: DT	Valoración
<p><i>Describir cómo fue realizada e interpretada la prueba índice:</i></p> <p>La prueba índice es el Termómetro de distrés (DT), que se aplica de manera heteroadministrada, incluida dentro de la entrevista diagnóstica. La interpretación de la prueba índice puede considerarse objetiva, dado que la puntuación del ítem es la respuesta directa aportada por el o la paciente.</p>		
<p><i>Preguntas orientativas:</i></p>		
<p>– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?</p>	<p>Sí</p>	
<p><b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b></p>		<p><input checked="" type="checkbox"/> <b>Bajo</b></p>
<p><b>DOMINIO 3: Prueba de referencia</b></p>		
<p><b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b></p>		<p><input checked="" type="checkbox"/> <b>Alto</b></p>
<p><i>Describir cómo fue realizada e interpretada la prueba de referencia:</i></p> <p>La prueba de referencia fue la Escala Hospitalaria de Ansiedad y Depresión (HADS). La HADS fue respondida de manera autoaplicada por el paciente en su domicilio, para posteriormente entregarla en el hospital. La interpretación de la prueba de referencia puede considerarse objetiva, dado que fue autoadministrada, y las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente. No obstante, el hecho de que la escala fuese respondida en el domicilio del paciente hace que no sea posible asegurar que fuese él o ella de manera individual quien haya respondido. Por otra parte, el ámbito de aplicación habitual de la HADS es el entorno sanitario (ambulatorio u hospitalario) y no el domiciliario, por lo que se desconoce si esto pudiera suponer algún sesgo en las respuestas de los y las pacientes y, por ende, en las puntuaciones de la HADS. La otra prueba de referencia utilizada fue una combinación de la HADS con una entrevista semiestructurada basada en los criterios diagnósticos DSM-IV, considerándose como casos solamente a los y las pacientes que se detectaron de modo convergente en ambas pruebas, es decir, quienes, teniendo una puntuación mayor o igual a 15 en la HADS, también se valorasen como casos mediante entrevista clínica. Llama la atención que el criterio haya sido combinado, cuando suele considerarse que una entrevista diagnóstica es la mejor prueba de referencia, por lo que podrían haberse contemplado únicamente los resultados de la entrevista diagnóstica como prueba de referencia. Se desconoce si el hecho de haber tenido en cuenta la convergencia de ambas pruebas podría comportar algún sesgo a la hora de determinar los casos.</p>		
<p><i>Preguntas orientativas:</i></p>		
<p>– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?</p>	<p>Incierto</p>	
<p>– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?</p>	<p>Incierto</p>	
<p><b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b></p>		<p><input checked="" type="checkbox"/> <b>Baja</b></p>
<p><b>DOMINIO 4: Flujo y tiempos</b></p>		
<p><b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b></p>		<p><input checked="" type="checkbox"/> <b>Incierto</b></p>
<p><i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i></p> <p>Se analizaron las características de 82 pacientes no participantes (quienes, aun cumpliendo los criterios de inclusión y exclusión, rechazaron participar en el estudio o no devolvieron los cuestionarios cubiertos). El 67,7 % eran mujeres. Se estudiaron de forma comparativa las características sociodemográficas de las personas participantes y no participantes, en términos de edad, estado civil, nivel educativo, diagnóstico oncológico, estadio, tiempo desde diagnóstico, antecedentes psicopatológicos y puntuación del Termómetro de distrés. Solo existen diferencias significativas en nivel educativo y en antecedentes psicopatológicos. Mientras que el nivel educativo no influyó en los niveles de malestar emocional, los antecedentes psicopatológicos sí se relacionaron con la presencia de malestar emocional.</p>		
<p><i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i></p> <p>El intervalo de tiempo entre la prueba índice (DT) y la de referencia es indeterminado, dado que la prueba índice se efectuó durante la entrevista diagnóstica una vez que se derivaron al programa, y la prueba de referencia (HADS) fue autoaplicada por parte de los propios pacientes en su domicilio, para luego entregarla en el hospital. Se desconoce cuánto tiempo transcurrió entre la aplicación de una y otra prueba, así como si entre ambas aplicaciones existe alguna intervención en este ámbito.</p>		
<p><i>Preguntas orientativas:</i></p>		
<p>¿hubo un intervalo apropiado entre la prueba índice y la de referencia?</p>	<p>Incierto</p>	
<p>¿recibieron todos los pacientes la misma prueba de referencia?</p>	<p>Sí</p>	
<p>¿se incorporaron todos los pacientes al análisis?</p>	<p>No</p>	
<p><b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Am- bulatoria. 2015; 18: 22-26.</p>		

Autoría/año: Martínez, 2015 (98)	Herramienta: DT	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<p><i>Describir el método de selección de pacientes:</i></p> <p>La población objeto de estudio son pacientes con cáncer en España. Los criterios de inclusión fueron: a) ser mayor de 18 años; b) tener un diagnóstico de cáncer; y c) ser capaz de proporcionar el consentimiento informado. Se invitó a participar a una muestra consecutiva de pacientes que visitaron algún departamento médico en la Fundación Instituto Valenciano de Oncología (FIVO). El 6 % de los y las pacientes que se abordaron (25 de 410) declinaron participar. La muestra final estuvo compuesta por 385 pacientes oncológicos. La edad media fue de 58,7 años (dt = 13,2). El 55 % eran hombres.</p>		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?	Sí	
– ¿se evitó un diseño de casos y controles?	Sí	
– ¿se evitaron exclusiones inapropiadas?	Sí	
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<p><i>Describir cómo fue realizada e interpretada la prueba índice:</i></p> <p>La prueba índice es el DT, y la combinación del DT con el IT. Ambas pruebas se aplicaron de manera autoinformada, como parte del conjunto de cuestionarios administrados. La interpretación de las pruebas índice puede considerarse objetiva, dado que fueron autoadministradas, y la puntuación del ítem es la respuesta directa aportada por el o la paciente.</p>		
<i>Preguntas orientativas:</i>		
– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?	Sí	
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<p><i>Describir cómo fue realizada e interpretada la prueba de referencia:</i></p> <p>La prueba de referencia es el BSI-18. Esta prueba se aplicó de manera autoinformada, como parte del conjunto de cuestionarios administrados. La interpretación de la prueba de referencia puede considerarse objetiva, dado que fue autoadministrada y las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente. Posteriormente, las puntuaciones totales se pasan a una escala T estandarizada. El punto de corte de la escala se explicita con anterioridad. Se aportan referencias de estudios previos que avalan las propiedades psicométricas satisfactorias de la prueba de referencia en población oncológica.</p>		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?	Sí	
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?	Sí	
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<p><i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i></p> <p>Hubo 25 participantes que rechazaron participar en el estudio (el 6 % de los y las pacientes que se abordaron). Las razones esgrimidas para ello fueron: no tener interés (45 %), estar demasiado cansados (28 %), tener mucha prisa (10 %) y otras razones (17 %). Se desconoce cuáles son las características de estas personas, pues no se analizaron.</p>		
<p><i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i></p> <p>Tanto la prueba índice como la prueba de referencia se administraron consecutivamente, ya que forman parte del conjunto de cuestionarios autoadministrados que ha de cubrir cada participante.</p>		

Autoría/año: Martínez, 2015 (98)	Herramienta: DT	Valoración
<i>Preguntas orientativas:</i>		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?		Sí
– ¿recibieron todos los pacientes la misma prueba de referencia?		Sí
– ¿se incorporaron todos los pacientes al análisis?		Sí
<b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.		

Autoría/año: Andreu, 2015 (35)	Herramienta: DT	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir el método de selección de pacientes:</i>		
La población objeto de estudio son pacientes adultos oncológicos hospitalizados con neoplasias hematológicas en España. Los criterios de selección fueron: a) tener un diagnóstico de neoplasia hematológica; b) no encontrarse en situación de aislamiento; y c) tener 18 años o más. La muestra inicial estaba formada por 100 pacientes consecutivos a quienes habían ingresado en el Servicio de Hematología del Hospital Universitario la Fe de Valencia entre febrero de 2013 y febrero de 2014. Un 19 % rechazó participar, por lo que la muestra final estuvo compuesta por 81 pacientes, de quienes un 53 % son varones. Todos los pacientes estaban ingresados en el hospital debido al propio tratamiento (90 %) o a complicaciones médicas (10 %).		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Sí
– ¿se evitó un diseño de casos y controles?		Sí
– ¿se evitaron exclusiones inapropiadas?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Bajo
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cómo fue realizada e interpretada la prueba índice:</i>		
La prueba índice es el DT, y la combinación del DT con el IT. No se especifica si las pruebas fueron autoadministradas o heteroadministradas. No se indica si las pruebas índice se realizaron antes o después de la prueba de referencia. La interpretación de las pruebas índice puede considerarse objetiva, dado que la puntuación del ítem es la respuesta directa aportada por el o la paciente.		
<i>Preguntas orientativas:</i>		
– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i>		
La prueba de referencia es el BSI-18. Se trata de un autoinforme compuesto por 18 ítems. No se menciona si la prueba de referencia se llevó a cabo antes o después de las pruebas índice. La interpretación de la prueba de referencia puede considerarse objetiva, dado que fue autoadministrada y las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente. Posteriormente, se pasan a una escala T estandarizada. El punto de corte de la escala se explicita con anterioridad. Se aportan referencias de estudios previos que avalan las propiedades psicométricas satisfactorias de la prueba de referencia en población oncológica.		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?		Sí
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		<input checked="" type="checkbox"/> Baja

Autoría/año: Andreu, 2015 (35)	Herramienta: DT	Valoración
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i> Hubo 19 pacientes que rechazaron participar en el estudio (el 19 % de la muestra inicial). Se desconoce cuáles son las características de estas personas, pues no se analizaron.		
<i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i> Se desconoce cuál fue el intervalo de tiempo existente entre la aplicación de las pruebas índice y la de referencia. Asimismo, se desconoce si hubo algún tipo de intervención entre la prueba índice y las de referencia, pero se intuye que se administraron el mismo día de manera consecutiva.		
<i>Preguntas orientativas:</i>		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?		Sí
– ¿recibieron todos los pacientes la misma prueba de referencia?		Sí
– ¿se incorporaron todos los pacientes al análisis?		Sí
<b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.		

Autoría/año: Andreu, 2019 (99)	Herramienta: DT	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir el método de selección de pacientes:</i> La población objeto de estudio son personas supervivientes españolas tras un tratamiento de cáncer de inicio en la edad adulta. Los criterios de inclusión fueron: a) tener 18 años o más; b) haber tenido un diagnóstico de cáncer, pero no padecer síntomas o signos de cáncer en ese momento; c) haber finalizado la cirugía, la quimioterapia o la radioterapia al menos 4 semanas antes del estudio; d) tener conocimientos de lengua española. Se invitó a participar a una muestra consecutiva inicial de 254 pacientes que acudían a la Fundación Instituto Valenciano de Oncología. El 93 % (236) dieron su consentimiento informado y completaron el conjunto de cuestionarios. La muestra final estaba formada por 236 participantes cuya media de edad era 58 años (dt = 13,4). El 65 % eran mujeres.		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Sí
– ¿se evitó un diseño de casos y controles?		Sí
– ¿se evitaron exclusiones inapropiadas?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Bajo
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<i>Describir cómo fue realizada e interpretada la prueba índice:</i> La prueba índice es el DT, que se aplicó de manera autoinformada, como parte del conjunto de cuestionarios administrados. La interpretación de la prueba índice puede considerarse objetiva, dado que fue autoadministrada, y la puntuación del ítem es la respuesta directa aportada por el o la paciente.		
<i>Preguntas orientativas:</i>		
– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo

Autoría/año: Andreu, 2019 (99)	Herramienta: DT	Valoración
<i>Describir cómo fue realizada e interpretada la prueba de referencia:</i>		
La prueba de referencia es el BSI-18, y se aplicó de manera autoinformada, como parte del conjunto de cuestionarios administrados. La interpretación de la prueba de referencia puede considerarse objetiva, dado que fue autoadministrada, y las puntuaciones de cada ítem se aplican según la respuesta directa aportada por el o la paciente. Posteriormente, las puntuaciones totales se pasan a una escala T estandarizada. El punto de corte de la escala se explicita con anterioridad. Se aportan referencias de estudios previos que avalan las propiedades psicométricas satisfactorias de la prueba de referencia en población oncológica.		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?		Sí
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		<input checked="" type="checkbox"/> <b>Baja</b>
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i>		
Hubo 19 pacientes que rechazaron participar en el estudio (el 7 % de la muestra inicial). Se desconoce cuáles son las características de estas personas, pues no se analizaron.		
<i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i>		
Se desconoce cuál fue el intervalo de tiempo existente entre la aplicación de las pruebas índice y la de referencia. Asimismo, se desconoce si hubo algún tipo de intervención entre la prueba índice y las de referencia, pero se intuye que se administraron el mismo día de manera consecutiva.		
<i>Preguntas orientativas:</i>		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?		Sí
– ¿recibieron todos los pacientes la misma prueba de referencia?		Sí
– ¿se incorporaron todos los pacientes al análisis?		Sí
<b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.		

Autoría/año: Limonero, 2012 (102)	Herramienta: DME	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<i>Describir el método de selección de pacientes:</i>		
La población objeto de estudio es la población oncológica-paliativa. Los criterios de inclusión fueron: a) pacientes con enfermedad oncológica avanzada a quienes habían hospitalizado; b) pacientes mayores de 18 años; c) facilitar el consentimiento informado. Se excluyeron los y las pacientes que no cumplieran estos criterios o que, aun cumpliéndolos, estaban en una situación clínica que impedía responder el cuestionario, presentaban deterioro cognitivo según la escala Pfeiffer (mayor o igual a 2 errores) o no eran capaces de mantener una conversación o comprender las preguntas del cuestionario. Del total de pacientes consecutivos que se analizaron, se excluyeron cuatro (3,6 %) por no cumplir alguno de los requisitos, de forma que la muestra final quedó constituida por 105 pacientes (50 hombres y 55 mujeres) con una media de edad de 67,3 ± 12,6 años.		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?		Sí
– ¿se evitó un diseño de casos y controles?		Sí
– ¿se evitaron exclusiones inapropiadas?		Sí
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Alto</b>

Autoría/año: Limonero, 2012 (102)	Herramienta: DME	Valoración
<p><i>Describir cómo fue realizada e interpretada la prueba índice:</i></p> <p>La prueba índice es el Cuestionario de Detección de Malestar Emocional (DME). El cuestionario fue administrado por personal médico o de enfermería durante los tres días posteriores al ingreso en la unidad de cuidados paliativos. El DME otorga una puntuación total (0 a 20) formada por la suma de las puntuaciones de las respuestas a las preguntas referentes a la valoración del estado de ánimo y a la de cómo se lleva la situación. La interpretación de la prueba índice puede considerarse objetiva, dado que la puntuación otorgada a los ítems coincide directamente con la respuesta aportada por el o la paciente. No queda claro por qué se termina seleccionando el punto de corte mayor o igual a 9, cuando, al analizar la curva ROC, el punto óptimo era 9,25 o 9,75.</p>		
<p><i>Preguntas orientativas:</i></p>		
<p>– ¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?</p>	<p>Sí</p>	
<p><b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b></p>		<p><input checked="" type="checkbox"/> Baja</p>
<p><b>DOMINIO 3: Prueba de referencia</b></p>		
<p><b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b></p>		<p><input checked="" type="checkbox"/> Bajo</p>
<p><i>Describir cómo fue realizada e interpretada la prueba de referencia:</i></p> <p>La prueba índice es la entrevista clínica semiestructurada, realizada de forma independiente por un psicólogo o una psicóloga con experiencia en cuidados paliativos, y evalúa a nivel global el malestar emocional de la persona enferma con una escala tipo Likert de 5 puntos (desde 1 = Ningún malestar; a 5 = Malestar emocional máximo). La entrevista tenía una duración que oscilaba entre 15 y 20 minutos, y la administró una psicóloga o un psicólogo durante los tres días posteriores al ingreso en la unidad de cuidados paliativos, en el intervalo de las dos horas siguientes después de haber administrado el DME por parte del personal médico o de enfermería. Asimismo, el profesional de la psicología administró el resto de los instrumentos (DT y HADS).</p>		
<p><i>Preguntas orientativas:</i></p>		
<p>– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?</p>	<p>Sí</p>	
<p>– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?</p>	<p>Sí</p>	
<p><b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b></p>		<p><input checked="" type="checkbox"/> Baja</p>
<p><b>DOMINIO 4: Flujo y tiempos</b></p>		
<p><b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b></p>		<p><input checked="" type="checkbox"/> Bajo</p>
<p><i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i></p> <p>Del total de pacientes consecutivos que se analizaron, se excluyeron cuatro (3,6 %) por no cumplir alguno de los requisitos, de forma que la muestra final quedó constituida por 105 pacientes. Se desconoce cuáles son las características de las personas no incluidas, ya que no se analizaron. En la conclusión, se menciona que se ha comparado la puntuación del DME con la valoración clínica, así como la del DT y, en algunos pacientes, con la de la HADS. Sin embargo, no se aportan datos de a cuántas personas se les ha administrado la HADS en concreto, ni se han analizado si existen diferencias significativas en las características de estas personas con respecto a quienes sí han respondido. Con todo, hemos tenido en cuenta la puntuación de la HADS principalmente para calcular la validez convergente.</p>		
<p><i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i></p> <p>El intervalo de tiempo transcurrido entre la administración de la prueba índice y la de referencia es de un máximo de dos horas. Aunque no se explicita, no parece que exista ninguna intervención en este ámbito entre ambas aplicaciones.</p>		
<p><i>Preguntas orientativas:</i></p>		
<p>– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?</p>	<p>Sí</p>	
<p>– ¿recibieron todos los pacientes la misma prueba de referencia?</p>	<p>Sí</p>	
<p>– ¿se incorporaron todos los pacientes al análisis?</p>	<p>Sí</p>	
<p><b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.</p>		

Autoría/año: Mateo-Ortega, 2019 (103)	Herramienta: ENP-E	Valoración
<b>DOMINIO 1: Selección de pacientes</b>		
<b>A. Riesgo de sesgo: ¿podría la selección de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Incierto
<p><i>Describir el método de selección de pacientes:</i></p> <p>La población objeto de estudio son pacientes al final de la vida. Los criterios de inclusión son: (1) diagnóstico de cáncer avanzado o terminal; (2) edad mayor o igual a 18; (3) consentimiento informado; y (4) estar recibiendo tratamiento de los equipos psicosociales y de cuidados paliativos. Los criterios de exclusión son: (1) presencia de condiciones que impedirían al paciente completar el cuestionario; (2) deterioro cognitivo (puntuación del test Pfeiffer &gt;3); o (3) negarse a dar el consentimiento informado. La muestra seleccionada estuvo formada por 150 pacientes que recibían tratamiento de las unidades de cuidados paliativos de los 19 hospitales participantes durante 3 meses en 2017: 21 pacientes formaron parte del estudio piloto y 129 pacientes formaron parte de la fase de estudio. La información de un paciente se descartó por estar incompleta. El 51 % eran mujeres. La edad media en hombres fue de 72 años (rango entre 33 y 93); la edad media en mujeres fue de 73 años (rango entre 43 y 93). No se explicita si se trata de una muestra consecutiva o aleatoria.</p>		
<i>Preguntas orientativas:</i>		
– ¿se incluyó una muestra consecutiva o aleatoria de pacientes?	Incierto	
– ¿se evitó un diseño de casos y controles?	Sí	
– ¿se evitaron exclusiones inapropiadas?	Sí	
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que los pacientes incluidos y ámbitos no coincidan con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 2: Prueba índice</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba índice haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Bajo
<p><i>Describir cómo fue realizada e interpretada la prueba índice:</i></p> <p>La prueba índice es el Instrumento de Evaluación de Necesidades Psicosociales y Espirituales del Enfermo al Final de la Vida (ENP-E). La sección A de la prueba consta de 13 ítems que preguntan al paciente sobre sentimientos, pensamientos y actividades que ha hecho recientemente. La escala de respuesta es tipo Likert de 5 puntos. La sección B consta de seis preguntas sobre el tipo de preocupaciones que tiene el o la paciente. La escala de respuesta es tipo Likert de 5 puntos. La sección C consta de seis preguntas sobre signos externos de malestar emocional detectados por quien entrevista. Esta prueba fue heteroadministrada por profesionales de la psicología (el 85,2 % de los casos), el trabajo social (el 11,4 % de los casos) o personal facultativo (el 3,4 %) durante los tres días posteriores a su ingreso hospitalario. A la hora de emplearla como prueba de detección precoz, solo se tiene en cuenta la puntuación total de la sección A. Además, algunos ítems (1, 2, 4, 10 y 12) discriminan mejor que otros (aunque no aportan datos concretos), por lo que se estimaron preguntas clave, de manera que las puntuaciones mayores o iguales a 4 en alguno de estos ítems se perciben como posibles signos de alerta de necesidades psicosociales y espirituales. La interpretación de la prueba índice puede considerarse objetiva, dado que la puntuación de cada escala es la respuesta directa señalada por el o la paciente.</p>		
<i>Preguntas orientativas:</i>		
¿se interpretaron los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?	Sí	
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la conducción o la interpretación de la prueba índice no coincida con la pregunta de la revisión?</b>		<input checked="" type="checkbox"/> Baja
<b>DOMINIO 3: Prueba de referencia</b>		
<b>A. Riesgo de sesgo: ¿podría la realización o la interpretación de la prueba de referencia haber introducido sesgos?</b>		<input checked="" type="checkbox"/> Incierto
<p><i>Describir cómo fue realizada e interpretada la prueba de referencia:</i></p> <p>La prueba de referencia es una combinación de los valores de varias pruebas: HADS (<math>\geq 17</math>), junto con la puntuación del ítem de calidad de vida del EORTC QLQ-C15-PAL (<math>\leq 3</math>), la puntuación del DT (<math>&gt; 5</math>) y el grado de sufrimiento (percepción a largo plazo). Esta combinación se construyó para indicar si la prueba de referencia en su conjunto es positiva o no. No queda claro cómo puntúa el grado de sufrimiento y a partir de qué puntuación se considera positivo. No se explicita si las pruebas de referencia se realizaron antes o después de la prueba índice. Se desconoce si se aplicaron en formato autoadministrado o heteroadministrado. Se intuye que tanto la prueba índice como las pruebas de referencia se llevaron a cabo el mismo día, pero tampoco se explicita como tal. La interpretación de la prueba de referencia puede considerarse objetiva, dado que la puntuación de cada ítem es la respuesta directa aportada por el o la paciente.</p>		
<i>Preguntas orientativas:</i>		
– ¿es probable que la prueba de referencia clasifique correctamente la condición blanco?	Incierto	
– ¿se interpretaron los resultados de la prueba de referencia sin conocimientos de los resultados de la prueba índice?	Sí	
<b>B. Preocupaciones con respecto a la aplicabilidad: ¿hay preocupación de que la condición blanco, clasificada por la prueba de referencia, difiera de la pregunta?</b>		<input checked="" type="checkbox"/> Baja

Autoría/año: Mateo-Ortega, 2019 (103)	Herramienta: ENP-E	Valoración
<b>DOMINIO 4: Flujo y tiempos</b>		
<b>A. Riesgo de sesgo: ¿podría el flujo de pacientes haber introducido sesgos?</b>		<input checked="" type="checkbox"/> <b>Bajo</b>
<p><i>Describir cualquier paciente que no haya recibido la prueba índice y/o de referencia o haya sido excluido de la tabla 2x2:</i> La información de un paciente se descartó por estar incompleta. Se desconoce cuáles son las características de ese paciente.</p> <p><i>Describir el intervalo de tiempo y cualquier intervención entre la prueba índice y la de referencia:</i> Aunque no se explicita, se entiende que la prueba índice y las pruebas de referencia se administraron consecutivamente en los tres días siguientes al ingreso hospitalario. No existe ninguna intervención entre la administración de la prueba índice y la de referencia.</p>		
Preguntas orientativas:		
– ¿hubo un intervalo apropiado entre la prueba índice y la de referencia?	Sí	
– ¿recibieron todos los pacientes la misma prueba de referencia?	Sí	
– ¿se incorporaron todos los pacientes al análisis?	Sí	
<p><b>Fuente:</b> Ciapponi, A. QUADAS-2: instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. Evidencia. Actualización en la Práctica Ambulatoria. 2015; 18: 22-26.</p>		

