



Centre universitari adscrit a la



**Universitat
Pompeu Fabra**
Barcelona

**Herramientas PROMs para la Evaluación del dolor y la
Funcionalidad en el Dolor de Cuello:
Una Revisión sistemática.**

Trabajo final de grado (Curso 2023-24)

Alumno: Adrian Alias Pallarés,

Tutor: Javier Picañol

10/05/2024

(08302)

Índice:

Índice de tablas y figuras.....	4
Glosario.....	4
1. Resumen.....	5
2. Introducción.....	6
2.1: Marco teórico respecto al Dolor Cervical.....	6
2.2: Telemedicina y uso de las PROMs.....	7
2.3: Propiedades psicométricas de una PROM.....	8
3. Justificación.....	9
3.2: Hipótesis.....	9
4. Objetivos del estudio.....	10
5. Metodos.....	11
5.1: Pregunta de investigación.....	11
5.2 Protocolo y registro.....	11
5.3 Criterios de elegibilidad.....	11
5.4 Búsqueda de la literatura científica.....	13
5.5 Evaluación de la calidad de los estudios y Riesgo de Sesgo.....	13
5.6 Puntos Finales y Análisis de Subgrupos.....	14
5.7 Extracción de los datos	14
6. Resultados.....	15
6.1 Selección de estudios.....	15
6.2 Características de estudios.....	16
6.3 Evaluación del sesgo de estudio.....	17
6.4 Evaluación de la calidad de la evidencia.....	18
6.5 Análisis variables de resultados.....	18
6.5.1. AEQ (Cuestionario de Evitación-Resistencia):.....	19
6.5.2. NPQ (Cuestionario de Northwich Park).....	20

6.5.3.NDI (Índice de discapacidad de cuello).....	20
6.5.4.COM(Core Outcome Measure).....	20
6.5.5.CF-PDI(Inventario de Dolor y Discapacidad Craneofacial).....	21
6.5.6. PCS-cat (Escala de catastrofización del dolor).....	21
6.5.7.SFI-SP (Índice funcional de la columna vertebral).....	21
6.5.8.OMPQ (Cuestionario Örebro de Dolor Musculoesquelético.).....	21
7. Discusión.....	22
7.3 Implicaciones en la practica clínica e investigación fututra.....	23
7.4 Limitaciones del estudio.....	23
8. Conclusión.....	23
9. Referencias bibliográficas.....	24
10. Anexos	29
10.1 Anexo 1(Ecuación de búsqueda).....	29
10.2 Anexo 2 (Prisma Checklist).....	30
10.3 Anexo 3 (Cosmin Risk of Bias check list).....	35
10.4 Anexo 4 (COSMIN guideline for Systematic reviews of PROMS).....	50
10.5Anexo 5.Evaluación calidad de estudio.....	51
10.6 Anexo 6: Evaluación cuestionario AEQ(artículo de Hasenbring).....	59
10.7 Anexo 7. Cuestionario NPQ (versión española).....	60
10.8 Anexo 8.Cuestionario NDI (versión española).....	61
10.9 Anexo 9. Tabla cuestionario COM.....	62
10.10 Anexo 10. Tabla cuestionario CF-PDI	63
1011 Anexo 11. Cuestionario PCS.....	64
10.12 Anexo 12. Cuestionario SFI-SP.....	65

Índice de tablas y figuras:

Tabla 1. Criterios de inclusión y exclusión.....	12
Figura 1. Diagrama de flujos PRISMA.....	15
Tabla 2. Características de los artículos incluidos.....	16
Tabla 3. Evaluación de sesgo según COSMIN Risk of Bias Assessment.....	17
Tabla 4. Resumen de la calidad de la Evidencia GRADE.....	18

Glosario

- **PROM:** Medidas de resultados informados por pacientes “Patient report outcome measure”
- **PRISMA:** Protocolo de Elementos para Informar Revisiones Sistemáticas.
- **COSMIN:** Consensus-based Standards for the selection of health Measurement INstruments.
- **PRÓSPERO:** Registro Internacional Prospectivo de Revisiones Sistemáticas.
- **GRADE:** Grading of Recommendations Assessment, Development and Evaluation.
- **PSFS:** Escala Funcional Específica del Paciente (Patient-Specific Functional Scale).
- **AEQ:** Cuestionario de Evitación del Miedo (Anxiety Avoidance Questionnaire en inglés).
- **NPAD:** Escala de Discapacidad y Dolor de Cuello (Neck Pain and Disability Scale).
- **CF-PDI:** Índice de Discapacidad Craneofacial (Craniofacial Pain Disability Index en inglés).
- **COM:** Medida de Resultados Centrales (Core Outcome Measure en inglés).
- **IDC:** Índice de Discapacidad Cervical (Cervical Disability Index en inglés).
- **NDI:** Índice de Discapacidad del Cuello (Neck Disability Index en inglés).
- **NPQ:** Cuestionario de Dolor Northwick (Northwick Pain Questionnaire en inglés).
- **OMPQ:** Cuestionario de Dolor Musculoesquelético de Örebro (Örebro Musculoskeletal Pain Questionnaire en inglés).
- **PCS:** Escala de Catastrofización del Dolor (Pain Catastrophizing Scale en inglés).
- **SFI-SP:** Índice Funcional de Salud de la Columna Vertebral (Spine Functional Index en inglés).

1. **Resumen:**

Este estudio se enfoca en una revisión sistemática que analiza la validez y fiabilidad de las PROMs en español para evaluar dolor y funcionalidad en adultos mayores de 18 años con patologías cervicales. Siguiendo el protocolo PRISMA y las directrices de COSMIN, se estableció un marco metodológico riguroso. Se realizaron búsquedas en bases de datos bibliográficas como PubMed, Web of Science, Cochrane Library y SportDISCUS, utilizando criterios de inclusión y exclusión predefinidos y desarrollando fórmulas de búsqueda específicas. De los 478 resultados iniciales, se excluyeron duplicados y se evaluaron minuciosamente 14 estudios, de los cuales se incluyeron 8 en la revisión. Todos contenían evaluaciones de la psicometría de las herramientas de inclusión. Se identificaron 8 herramientas (AEQ, NPQ, NDI, COM, CF-PDI, PCS-cat, SFI-SP y OMPQ), con índices de confiabilidad y validez óptimos y buenos en su psicometría. La revisión concluye reforzando la importancia del uso de las PROMs en el tratamiento del dolor y la funcionalidad de patologías cervicales, subrayando la necesidad de realizar estudios con mayor calidad de evidencia y menor riesgo de sesgo.

Palabras clave: Dolor de cuello y/o cervical, PROM (Patient report outcome measure), Psicometría, Fiabilidad y Validez.

Abstract:

This study focuses on a systematic review that analyzes the validity and reliability of PROMs in Spanish to assess pain and functionality in adults over 18 years old with cervical pathologies. Following the PRISMA protocol and COSMIN guidelines, a rigorous methodological framework was established. Searches were conducted in bibliographic databases such as PubMed, Web of Science, Cochrane Library, and SportDISCUS, using predefined inclusion and exclusion criteria and developing specific search formulas. Out of 478 initial results, duplicates were excluded, and 14 studies were meticulously evaluated, of which 8 were included in the review. All studies contained evaluations of the psychometrics of the inclusion tools. Eight tools were identified (AEQ, NPQ, NDI, COM, CF-PDI, PCS-cat, SFI-SP, and OMPQ), with optimal and good reliability and validity indices in their psychometrics. The review concludes by reinforcing the importance of using PROMs in the treatment of pain and functionality in cervical pathologies, emphasizing the need for studies with higher quality evidence and lower risk of bias.

Keywords: Neck and/or cervical pain, PROM (Patient-reported outcome measure), Psychometrics, Reliability, and Validity.

2. Introducción

2.1 Marco teórico respecto al Dolor Cervical

Es bien conocido al margen de la ciencia que el dolor de cuello o cervical es uno de los más comunes dentro de la población. Si atenemos al dolor, este es definido como una experiencia sensorial y emocional desagradable asociada a una lesión real o potencial o descrita en los términos de dicha lesión (Raja. et. al 2020). En este aspecto, el dolor de cuello es de las patologías que generan más discapacidad con una prevalencia que supera el 30% anual (Cohen S. P., 2015; Kim R. et Al 2018). Se determina que un alto porcentaje de casos conseguirá superar el periodo agudo de dolor, con o sin tratamiento, y a su vez se **observa** que el 50% de estos seguirán remitiendo algún grado de dolor a lo largo de su vida (Cohen S. P. 2015).

Lo cierto es que nos encontramos con múltiples factores psicosociales que pueden aumentar el riesgo de cronificación y promover un cuadro de dolor, conllevando una falta de funcionalidad. Diversas revisiones definen que los factores de riesgo psicosociales más significativos para desencadenar la experiencia dolorosa, fueron el estado de ánimo deprimido y el conflicto con el puesto laboral (kim.R. et Al 2018). Por otro lado se observa que el dolor de cuello se asocia a comorbilidades, antecedentes de dolor o un amplio historial clínico incluyendo malos hábitos, factores psicosociales y accidentes de tráfico (Hogg-Johnson, S. et Al 2008).

Otras investigaciones reflejan la importancia de una exhaustiva revisión de la historia clínica para discernir entre dolor neuropático y/o mecánico, con la identificación de potenciales causas que han podido preceder al dolor (Cohen S. P., 2015). Además, el paradigma actual de evaluación en fisioterapia debe contemplar de manera necesaria las banderas rojas, donde encontramos una variedad de lesiones de grado IV que hemos de tener presente. Patologías que precisan una derivación inmediata para poder abordarlas desde un enfoque biomédico y/o multidisciplinario (Bier J. D. et. Al al. 2018). Además, dentro del marco evaluativo del dolor cervical, el diagnóstico por imagen puede conllevar un valor importante (Bier,J.D et. Al 2018). Por otro lado, también contamos con pruebas clínicas que presentan distintos niveles de sensibilidad y especificidad (Delitto, A et Al 1995).

Referente al tratamiento, se evidencian tendencias favorables de cara a la educación del paciente, permanecer activo, terapia manual entre otras terapias físicas y estrategias cognitivas, sobreponiéndose a terapias inconcluyentes como el abordaje farmacológico y las terapias complementarias como la acupuntura, la electroterapia y la termoterapia (Corp, N. et. Al 2021).

Finalmente, teniendo en cuenta el avance científico-técnico que hemos vivido a lo largo del último siglo, han habido grandes avances en cuanto a la telemedicina y la telerehabilitación. De hecho, las revisiones sistemáticas y meta-análisis de los últimos años demuestran que son efectivas para reducir el dolor en trastornos musculoesqueléticos (Phuphanich, M.E et al, 2021). Sin embargo, hemos de tener en cuenta la necesidad de optimización, investigación e inclusión de las herramientas que tenemos disponibles para maximizar la eficacia de ello.

2.2 Telemedicina y uso de las Patient Report Outcome Measures

Entendiendo que las afecciones musculoesqueléticas conforman cerca del 30% de las atenciones médicas en algunos países de la Unión Europea, ha habido un auge en la investigación y en los planes de acción para atender este contexto. Desde un punto estratégico se comprende que el avance de la tecnología sugiere un cambio de paradigma en cuanto a la atención de diversas patologías y es en este momento que multitud de estudios se centran en la validez y eficacia de la medicina digital (Molina-Garcia, P. et al, 2023). Esta conlleva múltiples beneficios ya que las intervenciones ofrecen un alto alcance con un menor costo-efectivo y relativamente accesibles para una gran población de pacientes que abordan limitaciones de tiempo y recursos (Hewitt, S., 2020).

Las medidas de resultados informadas por el paciente (PROM, por sus siglas en inglés) validadas y fiables aparecen como una herramienta que encaja en la atención en línea, útil para pronosticar y llevar a cabo un seguimiento y tratamiento adecuado para la patología. La selección de esta debe estar relacionada con el contenido el cual busca y debe medir la solidez de sus propiedades psicométricas (Reeve, B. B., 2013).

Desde el constructo de interés del paciente, las PROMS aparecen como una medida capaz de ayudar a tomar conciencia del estado de salud general y/o específico de la patología. Son herramientas que observan constructos como por ejemplo los dominios afectivo-emocionales, sensorio-discriminativos y cognitivo-evaluativos del dolor, discapacidad, factores psicosociales y calidad de vida. Esto mejora la calidad de la atención en fisioterapia ya que fomenta la toma de decisiones informada entre paciente y fisioterapeuta. Las PROMs a su vez se convierten en una medida para individualizar los tratamientos entendiendo que estos aportan un reporte de datos del estado del cliente. De este modo, el registro y control de la herramienta permite monitorizar el progreso del tratamiento y la patología. A su vez, se convierten como herramienta objetiva de medición de las decisiones tomadas para el abordaje de la lesión.

Desde el punto de vista científico, las PROMs contribuyen a la investigación clínica como fuente de recopilación sistemática de datos (Feenstra, M., 2023). Es por este motivo que existen varios PROMs para evaluar a los pacientes con dolor de cuello (Reeve, B. B., 2013). Entre ellas, las que se utilizan con

más frecuencia son el Índice de discapacidad del cuello (NDI), el cuestionario para el dolor de cuello (NPQ), la escala funcional específica del paciente (PSFS) la escala de discapacidad y dolor de cuello (NPAD) y la de discapacidades del brazo, hombro y mano (Ramirez, M.M et Al. 2023). Actualmente estudios concluyen no ser necesaria la creación de nuevas herramientas en forma de cuestionarios sino que lo realmente importante subyace en la evaluación de su medición y de sus resultados (Schellingerhout, J. M. 2012).

2.3 Propiedades psicométricas de una PROM

La psicometría es fundamental para evaluar la calidad de las PROMs ya que la efectividad de estas medidas depende de la aplicación de principios rigurosos que garanticen su (1) confiabilidad y (2) validez (Reeve, B. B., 2013).

La confiabilidad en psicometría se refiere a la precisión con la que un conjunto de puntajes de pruebas mide lo que pretende medir (Ebel, 1977). En términos estadísticos, se define como la consistencia de las mediciones realizadas por una prueba o instrumento en diferentes circunstancias o momentos. Un aspecto crucial para asegurar la coherencia de las puntuaciones obtenidas. Dentro de la evaluación estandarizada, existen diferentes tipos de confiabilidad que son fundamentales:

- Test-Retest: Evalúa la estabilidad de las respuestas a lo largo del tiempo al administrar la prueba en ocasiones separadas.
- Forma Alterna: Analiza la consistencia entre dos formas similares de la misma prueba o instrumento para minimizar la influencia de la memoria o la práctica entre las administraciones.
- Consistencia Interna: Examina la estabilidad de las puntuaciones a través de los elementos que conforman una prueba, utilizando métodos como la confiabilidad dividida y la confiabilidad interítem.
- Confiabilidad entre Evaluadores: Mide la consistencia entre dos o más evaluadores al puntuar una prueba o al observar el comportamiento de un individuo.

La confiabilidad se evalúa comúnmente mediante un coeficiente de correlación, que varía de 0 a +1.0. Valores más altos indican mayor consistencia, pero la interpretación puede variar según el constructo medido.

Respecto a la validez, se considera una propiedad esencial de cualquier instrumento de medición, que asegura que el instrumento evalúe lo que se supone que debe evaluar. Se deben considerar diferentes tipos de validez, como:

- Validez de Contenido: Evalúa la adecuación y representatividad del contenido del instrumento en relación con el constructo que se está midiendo.
- Validez de Criterio: Examina la relación entre las puntuaciones de la prueba y un criterio externo que se considera una medida válida del mismo constructo.
- Validez de Constructo: Evalúa la medida en que las puntuaciones de la prueba se relacionan con teorías o conceptos específicos en el ámbito psicológico.

En conclusión, es esencial valorar la disponibilidad y actualidad de estos parámetros en las cuestiones que queremos revisar. En esta misma línea, dado que las PROMs podrían ser un herramienta clave en la valoración del dolor cervical pero precisan de una validez y confiabilidad evidenciadas, esta revisión sistemática pretende revisar los estudios que abordan propiedades psicométricas de los PROMs aplicados en el dolor cervical en lengua castellana

3 Justificación

Los problemas asociados a patologías cervicales y el impacto que este supone sobre la calidad de vida exige de una evaluación detallada, específica y confiable. Las PROMs, como una herramienta de valoración para los pacientes, posee un alto potencial en la telemedicina ofreciendo un futuro prometedor y costo-efectivo en la evaluación de patologías musculoesqueléticas. La problemática aparece cuando se da uso de las mismas sin tener un visión clara de la validez y fiabilidad que reportan. Por esta razón, en esta revisión sistemática se pretende dar una perspectiva en cuanto a herramientas PROM en lengua española para la evaluación de dolor y funcionalidad en patologías de cuello, analizando la validez y fiabilidad que reportan para poder generar evidencia que optimice la toma de decisiones. De esta forma se pretende potenciar una correcta adecuación de las herramientas resaltando tanto las limitaciones presentes como los *gaps* de estudio aún por abordar.

3.1 Hipótesis

Hipótesis de trabajo: Encontramos una existencia significativa de PROMs para el dolor cervical que muestran evidencia limitada en cuanto a ciertas propiedades psicométricas referentes a la fiabilidad y validez de estas. .

H0: Las herramientas PROMs para la evaluación del dolor y funcionalidad cervical en lenguaje español no presentan diferencias significativas en términos de validez y fiabilidad, indicando limitaciones en estas propiedades psicométricas.

H1: Las herramientas PROMs para la evaluación del dolor y funcionalidad cervical en lenguaje español presentan diferencias significativas en términos de validez y fiabilidad.

4. Objetivos del estudio

-Primario

+Identificar y evaluar aspectos psicométricos vinculados a la fiabilidad y validez de las PROMs para la evaluación del dolor y funcionalidad en pacientes con patología cervical en lengua española.

-Secundario

+Definir las PROMs en lengua española más investigadas en la evaluación del dolor y la funcionalidad en pacientes con patología cervical.

+Analizar las PROMs en lengua española relacionadas con el miedo-evitación en pacientes con patología cervical.

+Definir las PROMs relacionadas con el catastrofismo en pacientes con patologías cervical.

+Investigar las PROMs para evaluar la multidimensionalidad del dolor en pacientes con patología cervical.

+Explorar las PROMs para evaluar la calidad de vida en pacientes con patología cervical.

5. Métodos

5.1 Pregunta de investigación

¿Cuál es la validez y fiabilidad de las herramientas PROMs en español actualmente utilizadas para evaluar el dolor y la funcionalidad en pacientes con dolor de cuello?

5.2 Protocolo y registro

Esta revisión sistemática se ha llevado a cabo conforme las recomendaciones del Protocolo de elementos para informar de revisiones sistemáticas (PRISMA) con el fin de dar rigurosidad y confianza sobre la extracción de los datos. Cabe mencionar que también se ha considerado el uso de las pautas y el manual de COSMIN para revisiones sistemáticas encargadas de medir resultados provenientes de PROMs. Se hizo uso del *checklist* de PRISMA (Page MJ, et Al 2020) para desarrollar el protocolo de revisión (Anexo 2). De esta forma se facilitó la evaluación y la replicación de la revisión. De este modo, se ofrece un enfoque sistemático acorde con los estándares internacionales y las mejores prácticas actualizadas con el objetivo de contribuir de forma significativa en el conocimiento sobre las PROMs para el dolor y funcionalidad en patologías de cuello. De igual modo, el protocolo será registrado en el registro internacional prospectivo de revisiones sistemáticas (PRÓSPERO) en caso de querer ser publicado en el futuro para asegurar transparencia y rigurosidad.

5.3 Criterios de elegibilidad

La efectividad y relevancia de esta revisión sistemática dependen en gran medida de la precisión en la selección de los estudios. Por ello, se establecieron criterios de inclusión y exclusión detallados que guiaron el proceso. Se consideraron para su inclusión aquellos estudios que (1) se centraron en población adulta (mayores de 18 años) diagnosticados con patologías cervicales, ya sea como signo/síntoma o enfermedad principal. Esta especificidad aseguró que los resultados obtenidos estuviesen directamente relacionados con la población de interés. Asimismo, la revisión se enfocó en (2) investigaciones que utilicen las *Patient-Reported Outcome Measures* (PROMs) como herramientas de evaluación para medir el dolor o la funcionalidad en pacientes con patologías cervicales. Este enfoque permitió una evaluación detallada de la eficacia de estas medidas en el contexto específico de las afecciones cervicales. Se incluyeron únicamente (3) estudios que abordasen propiedades vinculadas a la validez y fiabilidad de las PROMs en patologías cervicales. Esta inclusión brindó una perspectiva más amplia y comprensiva de la utilidad de estas herramientas desde diversas dimensiones. (4) La revisión se limitó a estudios escritos en español o inglés para garantizar una interpretación coherente y

una síntesis eficaz de la información (McGoey, K. E., et.al 2010) y (5) únicamente se incluyeron estudios en población española para evaluar la utilización de las PROMs en lengua castellana.

Se excluyeron aquellos estudios que (1) no se centraron específicamente en pacientes con diagnóstico de patología cervical o dolor de cuello, asegurando así la relevancia directa a la temática de interés. (2) Investigaciones que no utilizaron las PROMs como herramientas de evaluación para mantener la coherencia en el análisis de estudios que aborden la eficacia de estas medidas. (3) Estudios que (3) no proporcionaron información sobre la psicometría de las PROMs, garantizando la integridad y robustez de la revisión. (4) Estudios escritos en idiomas no mencionados previamente y también aquellos cuyo estudio psicométrico fuese hecho sobre una herramienta no escrita en lenguas del territorio español, todo ello para mantener la consistencia y comprensión durante la revisión. Finalmente, se excluyeron estudios que (5) incluyesen a pacientes con grandes comorbilidades que pudiesen influir significativamente en la detección del dolor o la funcionalidad relacionada con la patología cervical, asegurando la homogeneidad en la población estudiada.

	Inclusión	Exclusión
Población	Estudios que observen a pacientes adultos (mayores de 18 años) diagnosticados con patologías cervicales como signo o síntoma principal.	Estudios que no se centren en pacientes con diagnóstico de patología cervical o dolor de cuello
Diseño de estudio	Estudios que utilicen las Patient-Reported Outcome Measures (PROMs) como herramientas de evaluación que midan el dolor o la funcionalidad de pacientes con patologías cervicales.	Investigaciones que no usen PROMs
Outcome	Se Incluyen estudios observacionales que incluyan validez y fiabilidad de las PROMs en Patologías cervicales	Estudios que no aparezcan datos de validez y fiabilidad
Idioma	Se seleccionan estudios en Español o inglés.	Idiomas anteriormente no expuestos y estudios que no estén vinculados con el uso en población española.
Tipo de estudio	Estudios que el título encaje con los objetivos del trabajo con preferencia para los estudios observacionales	Estudios que incluyan a pacientes con grandes comorbilidades que puedan afectar a la detección de dolor o funcionalidad como consecuencia de una patología cervical

Tabla 1. Criterios de inclusión y Exclusión

5.4 Estrategia de Búsqueda

La estrategia de búsqueda utilizada se basó en varios ítems clave, incluyendo la población de interés (personas con dolor cervical), los constructos de interés, las propiedades de medida (evaluación de la validez y fiabilidad de los instrumentos de medición del dolor cervical), la factibilidad (facilidad de uso de los instrumentos de medición) y los criterios de exclusión (estudios en animales y en otros idiomas). Para cada uno de estos ítems se desarrollaron términos de búsqueda específicos, utilizando los operadores booleanos AND y OR para combinarlos adecuadamente. La búsqueda se realizó inicialmente en PubMed y Cochrane Library, posteriormente se adaptó para otras bases de datos como Web of Science y Sportsdiscuss. Se generaron varias versiones de la ecuación de búsqueda, comenzando con una versión detallada y específica, pero al obtener resultados limitados, se simplificó para aumentar la sensibilidad y asegurar una mayor cobertura de estudios relevantes.

Las ecuaciones finales para PubMed y Web of Science se centraron en términos relacionados con el dolor cervical, la medición del dolor y la calidad de vida, con la inclusión de términos específicos para la validación de instrumentos de medición y la selección de estudios en humanos y en inglés o español. Los resultados obtenidos de estas búsquedas proporcionaron una base para la selección de estudios relevantes que cumplen con los criterios de inclusión establecidos para el protocolo de revisión sistemática. De hecho, la ecuación de búsqueda específica queda adjunta como material suplementario (Anexo 1).

5.5 Evaluación de la calidad de los estudios y el riesgo de sesgo

Las herramientas PROMs se evaluaron por Adrián Alias y en caso de debate o confusión fueron debatidas en consenso con el tutor utilizando la herramienta de COSMIN Risk Bias Checklist (Mokkink, L. B., et al., 2018). Los estudios se valoraron en cuanto a riesgo de sesgo siendo clasificados como “Muy bueno”, “adecuado”, “dudoso”, “inadecuado” o “no aplicable (NA)” en 10 dominios distintos: (1) Desarrollo de PROM, (2) Validez del contenido, (3) validez estructural, (4) coherencia interna, (5) validez transcultural\invariancia de medición, (6) confiabilidad, (7) error de medición, (8) validez del criterio, (9) pruebas de hipótesis para la validez de constructo y (10) capacidad de respuesta.

A su vez, se utilizó el sistema GRADE (Aguayo-Albasini, J. L., et al. 2014) para clasificar la calidad de los estudios. Este sistema se encarga de ayudar en la evaluación de la calidad de la evidencia en función de cinco dominios: (1) Diseño del estudio (analizando si el diseño es adecuado para responder a la pregunta de investigación), (2) la precisión (determinando si los resultados son fiables), (3) la aplicabilidad (observando si es aplicable a la población o intervención de interés, (4) la consistencia (revisando si son consistentes) y (5) el sesgo de publicación. En función de los resultados

de la evaluación de dichos dominios, la evidencia quedó clasificada en una de las siguientes categorías: “Alta”, “moderada” o “Baja calidad”.

5.6 Puntos Finales y Análisis de Subgrupos

Acorde a las guías COSMIN para Revisiones Sistemáticas de PROMs, la estructura de búsqueda la diferencia en 3 grandes bloques. El apartado A, donde se realizan investigaciones bibliográficas con diferentes pasos (formular el objetivo de la revisión, los criterios de elegibilidad, realizar una búsqueda bibliográfica y seleccionar resúmenes y artículos de texto completo). El apartado B, destinado a evaluar las propiedades de búsqueda (este apartado profundiza considerando la validez de contenido y la estructura interna: consistencia, validez y validez intercultural) y el resto de propiedades como la fiabilidad, medida de error y sensibilidad entre otros. Finalmente, el apartado C aborda la interpretabilidad y factibilidad, las recomendaciones formuladas y el informe sistemático (Detallado y esquematizado en Anexo 4).

Después de incluir los estudios y evaluar la calidad y el riesgo de sesgo, se analizó la estructura interna y las propiedades de medida de las PROMs incluidas en la revisión. Analizando de esta manera la fiabilidad y la validez de las herramientas psicométricas. Para ello, dependiendo la disponibilidad de parámetros, se analizó el coeficiente alfa de cronbach (consistencia interna), el coeficiente de correlación intraclass (fiabilidad intrasujeto) o kappa de cohen para fiabilidad inter-examinador entre otros. En cuanto a la validez se consideró la de contenido, la de constructo y la de criterio mediante una evaluación descriptiva.

5.7 Extracción de los datos

A.A.P. como autor independiente realizó la extracción de datos, los cuales organizó y clasificó en hojas de cálculos previamente preparadas (Microsoft Excel).

6. Resultados

6.1 Selección de Estudios

La estrategia de búsqueda en PubMed, Cochrane y Web of Science y SportDISCUSS reveló 478 resultados. Después de eliminar duplicados (n=96), y excluir estudios no relacionados mediante la revisión del título y/o resumen (n=368), 14 estudios se revisaron de manera completa para considerar detalladamente el posible cumplimiento de los criterios de inclusión. Finalmente, 8 estudios fueron incluidos en el análisis cualitativo. La estrategia de búsqueda del estudio PRISMA se presenta en la Figura 1.

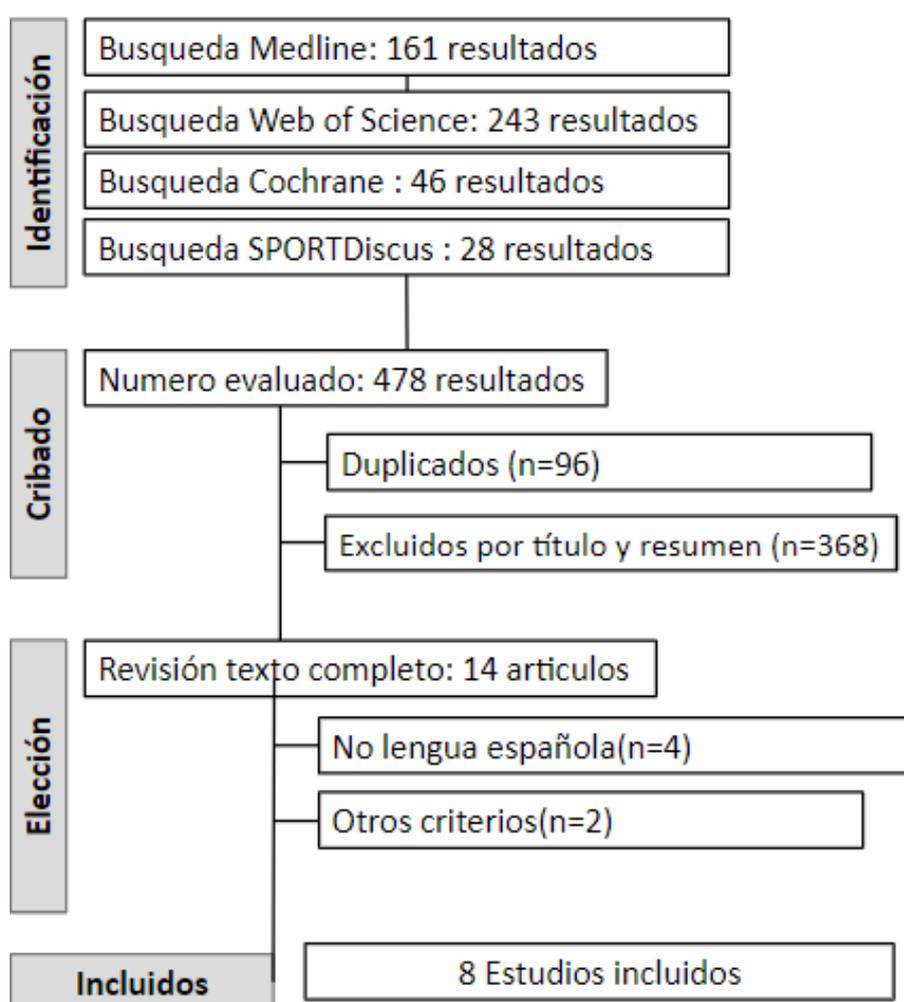


Figura 1. Diagrama de flujo PRISMA de selección y cribado del estudio

6.2 Características de los Estudios

En la presente revisión sistemática se consideraron los datos de un total de 3,983 participantes, donde una media del 68,42 % se trataba de sexo femenino. La edad media general fue de 44,30 años. La descripción del estudio acoge datos de 8 estudios observacionales (Tabla 2).

Estudio	Diseño del Estudio	Localidad	Tipo de población	N	F%	Edad media	Constructo	Test
Ruiz-Párraga et al. (2015)	OBSERVACIONAL	España (Málaga)	dolor musculoesquelético crónico	150	62	48,27	Confiabilidad y Validez	AEQ
M Kovacs, et al. (2008)	OBSERVACIONAL	España	Dolor de cuello	221	77,8	48,7	Confiabilidad y validez	NDI,NPQ, COM
La Touche, et al. (2014)	OBSERVACIONAL DESCRIPTIVO	España (Madrid)	Dolor y Discapacidad Craneofacial	192	45.1	45,7	Consistencia interna, Confiabilidad y validez	CF-PDI
Miró et al. (2007)	OBSERVACIONAL	España (Cataluña)	E1 dolor o pato crónica E2 Accidente con latigazo	146 +14 1	61	47,18+ 34,5	validez de criterio y fiabilidad	PCS
Cuesta-Vargas et al. (2014)	OBSERVACIONAL	España (Málaga)	Dolor columna +12 sem	226	58	45	validez constructo y confiabilidad y consistencia interna	SFI-SP
Ortega et al. (2007)	OBSERVACIONAL DESCRIPTIVO	España (Jaén)	Cervicalgia inespecífica o post traumática	48	81,2 5	38,75	validez constructo y confiabilidad y consistencia interna	NDI
Cuesta-Vargas et al. (2014)	OBSERVACIONAL DESCRIPTIVO	España (Málaga)	trastornos musculoesqueléticos crónicos	104	35,5 7	53,48	confiabilidad, estructura interna, consistencia interna y validez de criterio	OMPQ
Gabel et al. (2014)	OBSERVACIONAL DESCRIPTIVO	Internacional	Pacientes con latigazo cervical y dolor inespecífico cuello	175 5	68,7	37,2	Confiabilidad y validez interna	NDI
TOTAL	8 OBSERVACIONAL	7 España + 1 Int.	Relación PROMs + Cuello	398 3	68,4 2	44,30	Confiabilidad y validez interna, Consistencia interna, validez criterio	ALL on top

Tabla 2. Características de los estudios incluidos.

Detalles adicionales a la información en la tabla 2, En la tabla se muestran los 8 estudios seleccionados. En el apartado Estudio aparece la referencia, posteriormente se aprecia tipo de estudio, localidad tipo de población, (N) que representa el número de participantes en el estudio, F% hace referencia al porcentaje de participantes de género femenino. En constructo aparece el elemento buscado y seguido de test que es la PROM evaluada en el estudio. Abreviaciones (en orden alfabético): AEQ: Cuestionario de Evitación del Miedo (Anxiety Avoidance Questionnaire en inglés).CF-PDI: Índice de Discapacidad Craneofacial (Craniofacial Pain Disability Index en inglés).COM: Medida de Resultados Centrales (Core Outcome Measure en inglés).NDI: Índice de Discapacidad del Cuello (Neck Disability Index en inglés).NPQ: Cuestionario de Dolor Northwick (Northwick Pain Questionnaire en inglés). OMPQ: Cuestionario de Dolor Musculoesquelético de Örebro

(Örebro Musculoskeletal Pain Questionnaire en inglés).PCS: Escala de Catastrofización del Dolor (Pain Catastrophizing Scale en inglés). SFI-SP: Índice Funcional de Salud de la Columna Vertebral (Spine Functional Index en inglés).

La mayoría de estudios incluidos evaluaron la confiabilidad, la validez y la consistencia interna de las propiedades psicométricas de los cuestionarios entre los que podemos encontrar: AEQ, NDI, NPQ, COM, CF-PDI, PCS, SFI-SP, OMPQ entre otros (Tabla 2). Dada la presente heterogeneidad, los estudios no fueron cuantitativamente comparables.

6.3 Evaluación del sesgo

La interpretación del riesgo de sesgo de los estudios incluidos es crucial para garantizar la fiabilidad y validez de los datos extraídos. Para llevar a cabo esta evaluación, se empleó la herramienta COSMIN Risk of Bias Assessment. Esta herramienta evalúa 10 ítems (Tabla 3), pero dos de ellos fueron descartados (validez cultural y la validez de Constructo) al contener un resultado de “NA” en todas las celdas de la columna de los 8 estudios seleccionado.

Herramientas PROMs para la Evaluación del dolor y la Funcionalidad en el Dolor de Cuello: Una Revisión sistemática.								
Estudio	Diseño	Validez de contenido	Validez constructo	Consistencia interna	Confiabilidad	Error de medida	Validez de criterio	Capacidad respuesta
Ruiz-Párraga et al. (2015)	Adecuado	Adecuado	Adecuado	Adecuado	Adecuado	NA	NA	NA
M Kovacs ,et al.(2008)	Dudoso	Adecuado	Dudoso	Adecuado	Adecuado	NA	Dudoso	NA
La Touche,et al. (2014)	Muy bueno	Muy bueno	Adecuado	Adecuado	Adecuado	Adecuado	NA	Adecuado
Miró et al. (2007)	Adecuado	Adecuado	Muy bueno	Muy bueno	Muy bueno	NA	Adecuado	NA
Cuesta-Vargas et al. (2014) spine	Adecuado	NA	Adecuado	Muy bueno	Muy bueno	NA	Adecuado	Muy bueno
Ortega et al.(2007)	Adecuado	Dudoso	Muy bueno	Muy bueno	Muy bueno	NA	Muy bueno	Muy bueno
Cuesta-Vargas et al. (2014)	Adecuado	Adecuado	Adecuado	Muy bueno	Muy bueno	NA	Muy bueno	NA
Gabel et al. (2014)	Adecuado	Adecuado	Muy bueno	Muy bueno	Muy bueno	Adecuado	Adecuado	Adecuado

Tabla 3. Evaluación de sesgo según COSMIN Risk of Bias Assessment

Como se puede observar en la Tabla 3, se analizaron los restantes ítems para evaluar el riesgo de sesgo de las herramientas en relación con el estudio en cuestión, clasificando cada ítem según los criterios establecidos (Muy bueno, adecuado, dudoso, inadecuado o no evaluado).

6.4 Evaluación Calidad de la evidencia

Se utilizó el sistema de análisis GRADE (Grading of Recommendations, Assessment, Development and Evaluation) para evaluar de manera sistemática la calidad de la evidencia en base a varios dominios. En la Tabla 4 se puede observar la síntesis de la calidad GRADE de los estudios incluidos en este análisis, junto con la descripción del diseño del estudio y el instrumento de medición utilizado. De manera adicional, en el anexo 5 se describen los criterios de elección desglosados por estudio.

Contexto: No especificado		Bibliografía: GRADE System: Classification of Quality of Evidence and Strength of Recommendation						
Evaluación de calidad								
Estudio	Diseño del estudio/ instrumento de medición	Riesgo de sesgo	Inconsistencia	Indirectez	Imprecisión	Otro	Calidad GRADE	
Ruiz-Párraga et al., 2015	observacional transversal (p:150), Psicometría AEQ.	No detectable	Serio(-1)	No detectable	No detectable	Responde a lo esperado (+1)	Moderada	
M Kovacs ,et al., 2008	observacional transversal (P:221), Psicometría NDI, NPQ, COM.	Muy Serio(-2)	Serio(-1)	No detectable	Serio(-1)	Responde a lo esperado (+1) Gran efecto(+1)	Baja	
La Touche,et al., 2014	observacional prospectivo, transversal y descriptivo (P:192), CF-PDI.	Serio(-1)	No detectable	No detectable	Serio(-1)	Gran efecto(+1) Sin factores de confusió (+1) Responde a lo esperado(+1)	Moderada	
Miró et al. , 2007	Observacional, transversal y de cohorte prospectiva (P:426), Psicometría PCS.	Serio(-1)	No detectable	No detectable	No detectable	No se describe	Baja	
Cuesta-Vargas et al., 2014	Observacional prospectivo (P:226) Psicometría SFI-Sp.	No detectable	No detectable	No detectable	No detectable	Gran efecto(+1) Sin factores de confusió (+1) Responde a lo esperado(+2)	Moderada	
Ortega et al., 2007	Observacional y prospectivo (P:48) Psicometría IDC.	No detectable	No detectable	Serio(-1)	Muy Serio(-2)	Gran efecto(+2) Sin factores de confusió (+2) Responde a lo esperado(+2)	Baja	
Cuesta-Vargas et al., 2014	Observacional y prospectivo (P:104) Psicometría OMPQ,SF-12v2.	No detectable	No detectable	No detectable	Serio(-1)	Gran efecto(+2) Sin factores de confusió (+2) Responde a lo esperado(+2)	Moderada	
Gabel et al., 2014	Observacional (P:1755) Psicometría NDI.	Serio(-1)	No detectable	Serio(-1)	No detectable	Gran efecto(+2) Sin factores de confusió (+2) Responde a lo esperado(+2)	Moderada	

Tabla 4. Resumen de la calidad de la Evidencia GRADE

6.5 Análisis de las variables de resultados

El análisis de las variables de resultados da una visión detallada de diversas herramientas utilizadas en la evaluación del dolor para patologías de cuello y discapacidad asociada. A continuación,

se proporciona un breve análisis detallado del Cuestionario de Evitación-Resistencia (AEQ) hasta el Índice Funcional de la Columna Vertebral (SFI-SP), cada instrumento ofrece una perspectiva única sobre la experiencia del paciente y su funcionalidad. Los resultados de los estudios de validación y análisis psicométrico arrojan luz sobre la fiabilidad y validez de estas herramientas en contextos clínicos específicos.

6.5.1. AEQ (Cuestionario de Evitación-Resistencia):

El AEQ, conocido como Cuestionario de Evitación-Resistencia, es una herramienta diseñada para evaluar las respuestas de las personas ante el dolor, ya sea evitando actividades (FAR) o resistiendo a pesar del dolor (ER). Esta herramienta, basada en el estudio mencionado, mantiene su estructura original, compuesta por 20 ítems cuidadosamente seleccionados. Cada ítem se califica en una escala, que va desde 0 (totalmente en desacuerdo) hasta 6 (totalmente de acuerdo), lo que permite capturar la intensidad de la tendencia de evitación o resistencia. La puntuación total del cuestionario varía entre 0 y 120 puntos, reflejando así el grado en que una persona tiende a evitar o resistir las actividades a causa del dolor. En esta escala, puntuaciones más altas indican un mayor nivel de evitación o resistencia al dolor, mientras que puntuaciones más bajas sugieren una menor tendencia a evitar o resistir las actividades a pesar del dolor experimentado (Hasenbring ET.AL 2009).

Un artículo de calidad moderada incluido en la revisión aborda el AEQ. En él se buscaba analizar su estructura factorial y propiedades psicométricas en pacientes españoles con dolor crónico de espalda y cuello. Los resultados obtenidos indicaron que, si bien las escalas propuestas fueron en su mayoría respaldadas, la estructura factorial original del AEQ no pudo replicarse exactamente. Se encontró una diferenciación entre la depresión y la ansiedad, así como la distribución de los ítems cognitivos en las escalas de Catastrofización y Desamparo/Esperanza. Además, se identificaron cuatro escalas de comportamiento tras eliminar ciertos ítems, y se evidenció una alta consistencia interna y correlaciones bajas a moderadas entre las escalas.

Las escalas de evitación del dolor se relacionaron positivamente con la intensidad del dolor, la discapacidad funcional, la ansiedad relacionada con el dolor y otras creencias relacionadas con el miedo a evitar, mientras que las escalas de resistencia se asociaron negativamente con la discapacidad funcional y la ansiedad. Aunque este estudio posee limitaciones, los resultados sugieren que la AEQ-SP es una herramienta válida y confiable para identificar respuestas al dolor en pacientes españoles con dolor crónico de espalda y cuello, lo que podría mejorar la prevención y el tratamiento de estos pacientes.

6.5.2.NPQ (Cuestionario de Northwich Park)

El NPQ es un instrumento para proporcionar una evaluación de la discapacidad relacionada con el dolor de cuello, permitiendo a los profesionales de la salud comprender mejor el impacto funcional del dolor de cuello en la vida diaria de los pacientes. Utiliza una escala de puntuación desde el 0 al 100, donde el 0 representa ninguna limitación y el 100 representa la máxima limitación posible. Los pacientes deben marcar la opción que mejor describa su situación actual en cada dimensión evaluada. Las respuestas se ponderan y se calcula una puntuación total que refleja el grado de discapacidad relacionada con el dolor de cuello experimentada por el paciente. Un artículo con una calidad de evidencia baja determina que el NPQ no detecta un empeoramiento en la discapacidad y sugiere una mejora en el dolor en pacientes que niegan tal mejora. (M Kovacs, et al)

6.5.3.NDI (Índice de discapacidad de cuello)

El NDI es una herramienta esencial en la evaluación de dolor cervical, con más evidencia dentro de la bibliografía sobre las PROM de cuello. Esta tiene el diseño para evaluar la discapacidad asociada a dolores de cuello. Está compuesta por una serie de preguntas que abordan diferentes aspectos como la vida diaria y la función física. Cada pregunta se califica sobre 5 puntos y su puntuación total es de entre 0 y 50 puntos, donde mayor nivel de puntuación se asocian a mayor nivel de discapacidad.

Tras la revisión, el NDI se evalúa en 3 de los artículos midiendo y analizando sus propiedades psicométricas. El estudio Validación de una versión española del Neck Disability Index, con un GRADE de calidad Baja, validó la herramienta al español, con un alto coeficiente de Cronbach de 0.89, un coeficiente de correlación intraclase de 0,98 y correlaciones significativas con la escala visual analógica del dolor (EVA). En el artículo de Características psicométricas de la versión española de instrumentos para medir la discapacidad por dolor de cuello, se demostró un alto grado de consistencia interna y se demostró concordancia entre el NDI original y la versión española de este. El último estudio en el que se menciona aportó datos positivos en cuanto a la validez de constructo entre personas con latigazo cervical y otra población incrementando la confianza en su contenido.

6.5.4.COM (Core Outcome Measure)

El COM, se entiende como cuestionario de medida central después de una adaptación se diseñó para población con dolor cervical, esta PROM se encarga de medir la severidad del dolor, la función, el bienestar, la discapacidad, satisfacción y otros parámetros. Cada ítem del COM tiene 5 posibles respuestas. Las respuestas para los ítems 3 y 6 están ordenadas de peor a mejor, mientras que el resto están ordenadas de mejor a peor. El puntaje final se calcula como el promedio de los puntajes de cada ítem, y varía en un rango de 1.0 (mejor estado posible) a 5.0 (peor estado posible).

Tras la revisión, un artículo evaluó los datos psicométricos de la herramienta COM, este aportó resultados en cuanto a una alta consistencia con un alfa cronbach de 0,85, reportando una fiabilidad sólida del cuestionario. Además tuvo un tiempo corto de respuesta concluyendo como una herramienta rápida y confiable.

6.5.5.CF-PDI(Inventario de Dolor y Discapacidad Craneofacial)

El CF-PDI es un cuestionario autoadministrado diseñado para medir el dolor, la discapacidad y el estado funcional de las regiones mandibular y craneofacial en pacientes con dolor craneofacial crónico, Está compuesto de 22 items que acogen 4 aspectos principales: calidad de vida, estado funcional, evitación y dolor.

Tras la revisión se aprecia un estudio que lo menciona y expone su validación. Este demostró una alta consistencia interna con un coeficiente de cronbach de 0,88 y una alta fiabilidad con un Test-retest con correlación intraclase de 0,90. Se comparó este test con otros de dolor, discapacidad y calidad de vida y se demostró una asociación significativa.

6.5.6. PCS-cat (Escala de catastrofización del dolor)

Es un instrumento diseñado para evaluar la catastrofización frente al dolor, una escala que incluye 13 ítems evaluados con una puntuación de 5 puntos. Tras la revisión, un estudio evaluó su confiabilidad y validez en la traducción a la lengua catalana, la consistencia interna fue de 0,89 y aunque los resultados reportan una calidad dudosa y baja.

6.5.7.SFI-SP (Índice funcional de la columna vertebral)

El SFI es una herramienta PROM que evalúa toda la columna vertebral, la escala tiene una puntuación sobre 100 y se encarga de medir dolor, funcionalidad, discapacidad y calidad de vida de los encuestados. Tras la revisión un artículo proporciono datos psicométricos en cuanto fiabilidad, estructura factorial, consistencia interna y validez de criterio. El SFI-Sp demostró alta consistencia interna ($\alpha = 0,85$) y confiabilidad ($r = 0,96$). La estructura factorial fue unidimensional y apoyó la validez de constructo. La validez de criterio específico para la función fue alta con el RMQ ($r = 0,79$), moderada con el BADIX ($r = 0,59$) y baja con el NDI ($r = 0,46$).

6.5.8. OMPQ (Cuestionario Örebro de Dolor Musculoesquelético.)

El test Orebro es una herramienta que ayuda a identificar a los pacientes en riesgo a desarrollar problema musculoesqueléticos crónicos a largo plazo. Está compuesta de 25 preguntas, y una

calificación sobre 210, considerando que un puntaje menor a 105 puntos indica discapacidad baja, un puntaje entre 105 y 130 puntos indica discapacidad moderada, y un puntaje más alto indica discapacidad alta. Tras la revisión, en el artículo mostró datos psicométricos del OMPQ, la estructura interna mostró un análisis factorial que mostró un 75% varianza. Tras analizar la fiabilidad y consistencia interna, los valores de ICC estaban entre 0.853 y 1, mostrando una fiabilidad sobre el estudio alta. Hay una dimensión la de “coping” que mostró una fiabilidad baja.

7. Discusión

La presente revisión sistemática ha tratado de identificar y sintetizar las pruebas PROMs disponibles para la evaluación del dolor, la funcionalidad y otros parámetros como la sintomatología y calidad de vida, en patologías cervicales y del cuello. El uso de estas herramientas sin un criterio lógico y evidencia científica puede plantear problemas en la práctica clínica debido a la falta de confiabilidad y validez que algunas de estas aportan. Tras diseñar la pregunta de investigación con los parámetros PICO (Tabla 1), se llevó a cabo la búsqueda de literatura científica en PubMed, Cochrane Library, Web of Science y Sports Discuss, donde después de obtener 478 resultados, se identificaron 8 estudios observacionales que evaluaban la validez y confiabilidad de las PROMs en pacientes con patologías de cuello. Estos estudios abarcaron 8 herramientas PROMs, con un total de 3983 participantes con una edad media de 44 años, de los cuales el 68% eran mujeres. La evaluación del sesgo de los estudios se realizó con la herramienta COSMIN Risk of Bias Assessment para asegurar que las decisiones o análisis proporcionados estuvieran libres de influencias, prejuicios o suposiciones relacionadas con el cuestionario evaluado. Posteriormente, los estudios se analizaron con el análisis GRADE, y se reportaron 5 de ellos como de calidad Moderada y 3 como de Baja, destacando la herramienta NDI, seguida de otras que solo aparecieron reportadas una vez.

Atendiendo esencialmente al objetivo primario del estudio, se han identificado y evaluado múltiples PROMs que abarcan diferentes aspectos, desde funcionalidad hasta evaluación psicológica del dolor. Entre estas herramientas podemos mencionar el AEQ, NPQ, NDI, COM, CF-PDI, PCS-cat, SFI-SP y el OMPQ. Cada una ha sido analizada en términos de validez y confiabilidad, presentando los siguientes datos respectivamente: (Baja, Consistencia interna 0.83), (Alta, Cronbach 0.87), (CCI 0.98, Cronbach 0.89), (Validez consistente, Cronbach 0.85), (CCI 0.90, Cronbach 0.88), (Baja, Consistencia interna 0.89), (Baja, Consistencia interna 0.85), (>0.85, Alta).

Respecto a los objetivos secundarios del estudio, el AEQ destaca en el apartado de miedo-evitación en pacientes con patología cervical. En su versión española, ha mostrado evidencia moderada con datos medianamente adecuados. El estudio que se basa en la validación del AEQ aporta una diferenciación clara en la escala de comportamiento y correlaciones con la intensidad del dolor y la

discapacidad funcional. La PROM más citada y con evidencia robusta es el NDI, ampliamente reconocido y utilizado para el dolor y la discapacidad en patologías cervicales. En cuanto a las PROMs que identifican el catastrofismo en pacientes con patología cervical, se estudió la herramienta PCS-cat, que aportó una calidad de la evidencia baja, pero con datos consistentes en cuanto a confiabilidad y validez. Respecto a las herramientas que evalúan la multidimensionalidad, encontramos el CF-PDI, que, aunque se centra en la región craneofacial, proporciona un modelo de evaluación multidimensional del dolor que podría ser útil para obtener información relevante sobre el dolor cervical.

7.1 Implicaciones en Práctica clínica y investigación futura

Los resultados de esta revisión sistemática son importantes para respaldar la toma de decisiones en la práctica clínica y orientar futuras investigaciones. Los profesionales de la salud deben utilizar las PROMs más válidas y confiables para evaluar el dolor y la funcionalidad de los pacientes con lesiones cervicales o del cuello. Las PROMs presentadas pueden proporcionar información valiosa para la toma de decisiones clínicas, el seguimiento del progreso del paciente y el análisis de la eficacia de una intervención. Es crucial reconocer la necesidad de investigaciones adicionales; el objetivo debe ser validar los datos psicométricos de las PROMs existentes relacionadas con las patologías cervicales en el lenguaje español. También se necesitan estudios de alta calidad que aclaren la psicometría de las herramientas traducidas al español. Se requiere una investigación efectiva, precisa y sistemática para proporcionar datos confiables y válidos sobre las PROMs.

7.2 Limitaciones del estudio

La revisión presenta algunas limitaciones. En primer lugar, la búsqueda se realizó en artículos escritos en español o inglés y se seleccionaron solo aquellas PROMs cuyos datos psicométricos hubieran sido analizados en español. En segundo lugar, la calidad de la evidencia de los estudios fue moderada-baja, viéndose mermada la calidad final de la revisión sistemática. Además, debido a la heterogeneidad de los artículos encontrados, se hizo poco posible una óptima síntesis cuantitativa con meta-análisis.

8. Conclusiones

Esta revisión ofrece la importancia del uso de herramientas PROMs validadas y confiables en la evaluación de patologías de cuello y cervicales. Se observaron limitaciones en la calidad de la evidencia de los estudios analizados, lo que exige la necesidad de futuras investigaciones de calidad alta para validar los datos psicométricos de las herramientas en el contexto de dolor y funcionalidad cervical, especialmente en el idioma español. Es esencial que los profesionales de la salud seleccionen

las PROMs más evidenciadas para los pacientes a fin de mejorar la calidad de la atención clínica y el abordaje de patologías. Actualmente todas las PROMs descritas muestran valores de validez y confiabilidad medio-altos, aunque la mas analizada con un total de 3 estudios es el NDI.

9. Referencias bibliográficas

- *Aguayo-Albasini, J. L., Flores-Pastor, B., & Soria-Aledo, V. (2014). Sistema GRADE: clasificación de la calidad de la evidencia y graduación de la fuerza de la recomendación [GRADE system: classification of quality of evidence and strength of recommendation]. Cirugía Española, 92(2), 82–88. <https://doi.org/10.1016/j.ciresp.2013.08.002>*
- *Aguirre, M. V., Rodríguez, M. G., Clarett, M., Iribarne, J. I., Martínez, M., Battistotti, R., López de Arcaute, A. S., Adarves, R., & Orsini, E. (2013). Adaptación cultural y validación argentina del cuestionario Northwick park de dolor cervical en el ámbito hospitalario de la ciudad autónoma de Buenos Aires [Cultural adaptation and Argentine validation of the Northwick Park Neck Pain Questionnaire in the hospitals of the Autonomous City of Buenos Aires]. Revista de la Facultad de Ciencias Medicas (Cordoba, Argentina), 70(2), 76–82.*
- *Andrade Ortega, J. A., Delgado Martínez, A. D., & Almécija Ruiz, R. (2008). Validación de una versión española del Índice de Discapacidad Cervical [Validation of a Spanish version of the Neck Disability Index]. Medicina clinica, 130(3), 85–89. <https://doi.org/10.1157/13115352>*
- *Barbero, M., Moresi, F., Leoni, D., Gatti, R., Egloff, M., & Falla, D. (2015). Test-retest reliability of pain extent and pain location using a novel method for pain drawing analysis. European journal of pain (London, England), 19(8), 1129–1138. <https://doi.org/10.1002/ejp.636>*
- *Bier, J. D., Scholten-Peeters, W. G. M., Staal, J. B., Pool, J., van Tulder, M. W., Beekman, E., Knoop, J., Meerhoff, G., & Verhagen, A. P. (2018). Clinical Practice Guideline for Physical Therapy Assessment and Treatment in Patients With Nonspecific Neck Pain. Physical Therapy, 98(3), 162–171. <https://doi.org/10.1093/ptj/pzx118>*

- Cohen, S. P. (2015). *Epidemiology, diagnosis, and treatment of neck pain*. *Mayo Clinic Proceedings*, 90(2), 284–299. <https://doi.org/10.1016/j.mayocp.2014.09.008>
- Corp, N., Mansell, G., Stynes, S., Wynne-Jones, G., Morsø, L., Hill, J. C., & van der Windt, D. A. (2021). *Evidence-based treatment recommendations for neck and low back pain across Europe: A systematic review of guidelines*. *European Journal of Pain (London, England)*, 25(2), 275–295. <https://doi.org/10.1002/ejp.1679>
- Corral, Y. (2009). *Validez y confiabilidad de los instrumentos de investigación para la recolección de datos*. *Revista Ciencias de la Educación*, 19, 228–247.
- Cuesta-Vargas, A. I., & González-Sánchez, M. (2014). *Spanish version of the screening Örebro musculoskeletal pain questionnaire: a cross-cultural adaptation and validation*. *Health and quality of life outcomes*, 12, 157. <https://doi.org/10.1186/s12955-014-0157-5>
- Cuesta-Vargas, A. I., & Gabel, C. P. (2014). *Validation of a Spanish version of the Spine Functional Index*. *Health and quality of life outcomes*, 12, 96. <https://doi.org/10.1186/1477-7525-12-96>
- Delitto, A., & Snyder-Mackler, L. (1995). *The diagnostic process: examples in orthopedic physical therapy*. *Physical Therapy*, 75(3), 203–211. <https://doi.org/10.1093/ptj/75.3.203>
- Ebel, R. L. (1977). *Fundamentos de la medición educacional*. Guadalupe.
- Feenstra, M. M., Kidholm, K., & Rudnicki, M. (2023). *Use of patient reported outcome measures in daily clinical practice - A way to optimize treatment of patients?.* *Acta Obstetricia et Gynecologica Scandinavica*, 102(4), 404–405. <https://doi.org/10.1111/aogs.14501>
- Gabel, C. P., Cuesta-Vargas, A., Barr, S., Black, S. W., Osborne, J. W., & Melloh, M. (2016). *Confirmatory factor analysis of the neck disability index, comparing patients with whiplash associated disorders to a control group with non-specific neck pain*. *European spine journal : official publication of the European Spine Society, the*

European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society, 25(7), 2078–2086. <https://doi.org/10.1007/s00586-016-4543-z>

- Hasenbring, M. I., Hallner, D., & Rusu, A. C. (2009). Fear-avoidance- and endurance-related responses to pain: development and validation of the Avoidance-Endurance Questionnaire (AEQ). *European journal of pain (London, England)*, 13(6), 620–628. <https://doi.org/10.1016/j.ejpain.2008.11.001>
- Hewitt, S., Sephton, R., & Yeowell, G. (2020). The Effectiveness of Digital Health Interventions in the Management of Musculoskeletal Conditions: Systematic Literature Review. *Journal of Medical Internet Research*, 22(6), e15617. <https://doi.org/10.2196/15617>
- Hogg-Johnson, S., van der Velde, G., Carroll, L. J., Holm, L. W., Cassidy, J. D., Guzman, J., Côté, P., Haldeman, S., Ammendolia, C., Carragee, E., Hurwitz, E., Nordin, M., Peloso, P., & Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders (2008). The burden and determinants of neck pain in the general population: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *Spine*, 33(4 Suppl), S39–S51. <https://doi.org/10.1097/BRS.0b013e31816454c8>
- Kim, R., Wiest, C., Clark, K., Cook, C., & Horn, M. (2018). Identifying risk factors for first-episode neck pain: A systematic review. *Musculoskeletal Science & Practice*, 33, 77–83. <https://doi.org/10.1016/j.msksp.2017.11.007>
- Kovacs, F. M., Bagó, J., Royuela, A., Seco, J., Giménez, S., Muriel, A., Abreira, V., Martín, J. L., Peña, J. L., Gestoso, M., Mufraggi, N., Núñez, M., Corcoll, J., Gómez-Ochoa, I., Ramírez, M. J., Calvo, E., Castillo, M. D., Martí, D., Fuster, S., Fernández, C., ... Cano, A. (2008). Psychometric characteristics of the Spanish version of instruments to measure neck pain disability. *BMC musculoskeletal disorders*, 9, 42. <https://doi.org/10.1186/1471-2474-9-42>
- La Touche, R., Pardo-Montero, J., Gil-Martínez, A., Paris-Aleman, A., Angulo-Díaz-Parreño, S., Suárez-Falcón, J. C., Lara-Lara, M., & Fernández-Carnero, J. (2014).

Craniofacial pain and disability inventory (CF-PDI): development and psychometric validation of a new questionnaire. Pain physician, 17(1), 95–108.

- *McGoey, K. E., Cowan, R. J., Rumrill, P. P., & LaVogue, C. (2010). Understanding the psychometric properties of reliability and validity in assessment. Work (Reading, Mass.), 36(1), 105–111. <https://doi.org/10.3233/WOR-2010-1012>*
- *Miró, J., Nieto, R., & Huguet, A. (2008). The Catalan version of the Pain Catastrophizing Scale: a useful instrument to assess catastrophic thinking in whiplash patients. The journal of pain, 9(5), 397–406. <https://doi.org/10.1016/j.jpain.2007.12.004>*
- *Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Quality of Life Research, 27(5), 1171–1179. <https://doi.org/10.1007/s11136-017-1765-4>*
- *Neck Pain and Its Associated Disorders. Spine, 33(4 Suppl), S101–S122. <https://doi.org/10.1097/BRS.0b013e3181644ae8>*
- *Nordin, M., Carragee, E. J., Hogg-Johnson, S., Weiner, S. S., Hurwitz, E. L., Peloso, P. M., Guzman, J., van der Velde, G., Carroll, L. J., Holm, L. W., Côté, P., Cassidy, J. D., Haldeman, S., & Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders (2008). Assessment of neck pain and its associated disorders: results of the Bone and Joint Decade 2000-2010 Task Force on*
- *Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, 372, n71. doi: 10.1136/bmj.n71*
- *Pérez Fuentes, J. (2020). Versión actualizada de la definición de dolor de la IASP: un paso adelante o un paso atrás. Revista de la Sociedad Española del Dolor, 27(4), 232–233. Epub 13 de octubre de 2020. <https://dx.doi.org/10.20986/resed.2020.3839/2020>*

- Reeve, B. B., Wyrwich, K. W., Wu, A. W., Velikova, G., Terwee, C. B., Snyder, C. F., Schwartz, C., Revicki, D. A., Moinpour, C. M., McLeod, L. D., Lyons, J. C., Lenderking, W. R., Hinds, P. S., Hays, R. D., Greenhalgh, J., Gershon, R., Feeny, D., Fayers, P. M., Cella, D., Brundage, M., ... Butt, Z. (2013). *ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research*. *Quality of Life Research*, 22(8), 1889–1905. <https://doi.org/10.1007/s11136-012-0344-y>
- Rodriguez, A., Edelen, M. O., Herman, P. M., & Hays, R. D. (2022). *Unpacking the impact of chronic pain as measured by the impact stratification score*. *BMC musculoskeletal disorders*, 23(1), 884. <https://doi.org/10.1186/s12891-022-05834-4>
- Ruiz-Párraga, G. T., López-Martínez, A. E., Rusu, A. C., & Hasenbring, M. I. (2015). *Spanish Version of the Avoidance-Endurance Questionnaire: Factor Structure and Psychometric Properties*. *The Spanish journal of psychology*, 18, E88. <https://doi.org/10.1017/sjp.2015.89>
- Schellingerhout, J. M., Verhagen, A. P., Heymans, M. W., Koes, B. W., de Vet, H. C., & Terwee, C. B. (2012). *Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review*. *Quality of Life Research*, 21(4), 659–670. <https://doi.org/10.1007/s11136-011-9965-9>
- Thoomes-de Graaf, M., Fernández-De-Las-Peñas, C., & Cleland, J. A. (2020). *The content and construct validity of the modified patient specific functional scale (PSFS 2.0) in individuals with neck pain*. *The Journal of manual & manipulative therapy*, 28(1), 49–59. <https://doi.org/10.1080/10669817.2019.1616394>
- Young, I. A., Dunning, J., Butts, R., Cleland, J. A., & Fernández-de-Las-Peñas, C. (2019). *Psychometric properties of the Numeric Pain Rating Scale and Neck Disability Index in patients with cervicogenic headache*. *Cephalalgia : an international journal of headache*, 39(1), 44–51. <https://doi.org/10.1177/0333102418772584>

10. Anexos:

Anexo 1: Ecuación de búsqueda.

Primero se organizó los distintos ítems que debían aparecer en nuestra ecuación. En este caso, acorde a otras revisiones y los COSMIN guidelines, se diferencié a la Población, el Constructo de Interés, las Propiedades de Medida, la Factibilidad y los Criterios de exclusión.

Para el constructo de interés, las propiedades de medida y la factibilidad se han utilizado filtros ya utilizados y evidenciados anteriormente en otros estudios. Para los criterios de exclusión y la población me he basado en nuestro PICOTT framework.

Se adapta a pubmed, y junta los diferentes ítems con los operadores booleanos de “AND”. A posterior, como se encuentran algunos problemas de compatibilidad, se generan las ecuaciones mejor adaptadas para pubmed. Sin embargo, dado que los resultados dieron demasiado bajos, imposibilitando una buena búsqueda y selección de estudios... Se generó una simplificación, casi masiva, de los distintos ítems.

Resultando la siguiente ecuación.

Pubmed: ("neck pain" OR "cervical pain") AND ("pain measurement" OR "PROMs" OR "pain assessment" OR "patient reported outcome measures") AND ("Validation Studies" OR "Psychometrics" OR "Outcome Assessment, Health Care") AND ("humans"[MeSH Terms] AND ("english"[Language] OR "spanish"[Language]))

Cómo la búsqueda se ha de hacer mínimo en 3 bases de datos, se desarrolló una adaptación para Web of Science. Como la ecuación anterior no dió buenos resultados, se simplificó un poco más dejando este proceso:

Pubmed: ("neck pain" OR "cervical pain") AND ("pain" OR "pain assessment" OR "outcome measures") AND LANGUAGE: (English OR Spanish)

Resultados totales: 161

Web of Science: ("neck pain" OR "cervical pain") AND ("pain" OR "pain assessment" OR "outcome measures") AND LANGUAGE: (English OR Spanish)

Resultados totales: 243

A posterior se extrajo la ecuación de pubmed y se trató de adaptar al buscador de Cochrane dando la búsqueda reducida siguiente:

Adaptación a Cochrane: ("neck pain" OR "cervical pain") AND ("pain measurement" OR "PROMs" OR "pain assessment" OR "patient reported outcome measures") AND ("Validation Studies" OR "Psychometrics" OR "Outcome Assessment, Health Care")

Resultados totales: 46

Adaptación SportDiscuss: ("neck pain" OR "cervical pain") AND ("pain measurement" OR "PROMs" OR "pain assessment" OR "patient reported outcome measures") AND ("Validation Studies" OR "Psychometrics" OR "Outcome Assessment, Health Care")

Resultados totales: 28

Anexo 2: PRISMA checklist:

Tabla 1

Lista de verificación PRISMA 2020

Sección/tema	Ítem n.º	Ítem de la lista de verificación	Localización del ítem en la publicación
TÍTULO			
Título	1	Identifique la publicación como una revisión sistemática.	
RESUMEN			
Resumen estructurado	2	Vea la lista de verificación para resúmenes estructurados de la declaración PRISMA 2020 (tabla 2).	
INTRODUCCIÓN			
Justificación	3	Describa la justificación de la revisión en el contexto del conocimiento existente.	
Objetivos	4	Proporcione una declaración explícita de los objetivos o las preguntas que aborda la revisión.	
MÉTODOS			
Criterios de elegibilidad	5	Especifique los criterios de inclusión y exclusión de la revisión y cómo se agruparon los estudios para la síntesis.	
Fuentes de información	6	Especifique todas las bases de datos, registros, sitios web, organizaciones, listas de referencias y otros recursos de búsqueda o consulta para identificar los estudios. Especifique la fecha en la que cada recurso se buscó o consultó por última vez.	
Estrategia de búsqueda	7	Presente las estrategias de búsqueda completas de todas las bases de datos, registros y sitios web, incluyendo cualquier filtro y los límites utilizados.	
Proceso de selección de los estudios	8	Especifique los métodos utilizados para decidir si un estudio cumple con los criterios de inclusión de la revisión, incluyendo cuántos autores de la revisión cribaron cada registro y cada publicación recuperada, si trabajaron de manera independiente y, si procede, los detalles de las herramientas de automatización utilizadas en el proceso.	
Proceso de extracción de los datos	9	Indique los métodos utilizados para extraer los datos de los informes o publicaciones, incluyendo cuántos revisores recopilaron datos de cada publicación, si trabajaron de manera independiente, los procesos para obtener o confirmar los datos por parte de los investigadores del estudio y, si procede, los detalles de las herramientas de automatización utilizadas en el proceso.	

Lista de los datos	10a	Enumere y defina todos los desenlaces para los que se buscaron los datos. Especifique si se buscaron todos los resultados compatibles con cada dominio del desenlace (por ejemplo, para todas las escalas de medida, puntos temporales, análisis) y, de no ser así, los métodos utilizados para decidir los resultados que se debían recoger.
	10b	Enumere y defina todas las demás variables para las que se buscaron datos (por ejemplo, características de los participantes y de la intervención, fuentes de financiación). Describa todos los supuestos formulados sobre cualquier información ausente (<i>missing</i>) o incierta.
Evaluación del riesgo de sesgo de los estudios individuales	11	Especifique los métodos utilizados para evaluar el riesgo de sesgo de los estudios incluidos, incluyendo detalles de las herramientas utilizadas, cuántos autores de la revisión evaluaron cada estudio y si trabajaron de manera independiente y, si procede, los detalles de las herramientas de automatización utilizadas en el proceso.
Medidas del efecto	12	Especifique, para cada desenlace, las medidas del efecto (por ejemplo, razón de riesgos, diferencia de medias) utilizadas en la síntesis o presentación de los resultados.
Métodos de síntesis	13a	Describa el proceso utilizado para decidir qué estudios eran elegibles para cada síntesis (por ejemplo, tabulando las características de los estudios de intervención y comparándolas con los grupos previstos para cada síntesis (ítem n.8.5)).
	13b	Describa cualquier método requerido para preparar los datos para su presentación o síntesis, tales como el manejo de los datos perdidos en los estadísticos de resumen o las conversiones de datos.
	13c	Describa los métodos utilizados para tabular o presentar visualmente los resultados de los estudios individuales y su síntesis.
	13d	Describa los métodos utilizados para sintetizar los resultados y justifique sus elecciones. Si se ha realizado un metaanálisis, describa los modelos, los métodos para identificar la presencia y el alcance de la heterogeneidad estadística, y los programas informáticos utilizados.
	13e	Describa los métodos utilizados para explorar las posibles causas de heterogeneidad entre los resultados de los estudios (por ejemplo, análisis de subgrupos, metarregresión).
	13f	Describa los análisis de sensibilidad que se hayan realizado para evaluar la robustez de los resultados de la síntesis.

Tabla 1 (Continuación) Lista de verificación

PRISMA 2020

Sección/tema	Ítem n.º	Localización del ítem en la publicación	Ítem de la lista de verificación
Evaluación del sesgo en la publicación	14		Describa los métodos utilizados para evaluar el riesgo de sesgo debido a resultados faltantes en una síntesis (derivados de los sesgos en las publicaciones).
Evaluación de la certeza de la evidencia	15		Describa los métodos utilizados para evaluar la certeza (o confianza) en el cuerpo de la evidencia para cada desenlace.
RESULTADOS			
Selección de los estudios	16a		Describa los resultados de los procesos de búsqueda y selección, desde el número de registros identificados en la búsqueda hasta el número de estudios incluidos en la revisión, idealmente utilizando un diagrama de flujo (ver figura 1).
	16b		Cite los estudios que aparentemente cumplen con los criterios de inclusión, pero que fueron excluidos, y explique por qué fueron excluidos.
Características de los estudios	17		Cite cada estudio incluido y presente sus características.
Riesgo de sesgo de los estudios individuales	18		Presente las evaluaciones del riesgo de sesgo para cada uno de los estudios incluidos.
Resultados de los estudios individuales	19		Presente, para todos los desenlaces y para cada estudio: a) los estadísticos de resumen para cada grupo (si procede) y b) la estimación del efecto y su precisión (por ejemplo, intervalo de credibilidad o de confianza), idealmente utilizando tablas estructuradas o gráficos.
Resultados de la síntesis	20a		Para cada síntesis, resume brevemente las características y el riesgo de sesgo entre los estudios contribuyentes.
	20b		Presente los resultados de todas las síntesis estadísticas realizadas. Si se ha realizado un metanálisis, presente para cada uno de ellos el estimador de resumen y su precisión (por ejemplo, intervalo de credibilidad o de confianza) y las medidas de heterogeneidad estadística. Si se comparan grupos, describa la dirección del efecto.
	20c		Presente los resultados de todas las investigaciones sobre las posibles causas de heterogeneidad entre los resultados de los estudios.
	20d		Presente los resultados de todos los análisis de sensibilidad realizados para evaluar la robustez de los resultados sintetizados.
Sesgos en la publicación	21		Presente las evaluaciones del riesgo de sesgo debido a resultados faltantes (derivados de los sesgos de en las publicaciones) para cada síntesis evaluada.
Certeza de la evidencia	22		Presente las evaluaciones de la certeza (o confianza) en el cuerpo de la evidencia para cada desenlace evaluado.
DISCUSIÓN			
Discusión	23a		Proporcione una interpretación general de los resultados en el contexto de otras evidencias.
	23b		Argumente las limitaciones de la evidencia incluida en la revisión.
	23c		Argumente las limitaciones de los procesos de revisión utilizados.
	23d		Argumente las implicaciones de los resultados para la práctica, las políticas y las futuras investigaciones.
OTRA INFORMACIÓN			
Registro y protocolo	24a		Proporcione la información del registro de la revisión, incluyendo el nombre y el número de registro, o declare que la revisión no ha sido registrada.

	24b	Indique do´nde se puede acceder al protocolo, o declare que no se ha redactado ningun protocolo.
	24c	Describa y explique cualquier enmienda a la informaci3n proporcionada en el registro o en el protocolo.
Financiaci3n	25	Describa las fuentes de apoyo financiero o no financiero para la revisi3n y el papel de los financiadores o patrocinadores en la revisi3n.
Conflicto de intereses	26	Declare los conflictos de intereses de los autores de la revisi3n.
Disponibilidad de datos, co´digos y otros materiales	27	Especifique que´ elementos de los que se indican a continuaci3n est3n disponibles al p´blico y do´nde se pueden encontrar: plantillas de formularios de extracci3n de datos, datos extra´dos de los estudios incluidos, datos

Tabla 2

Lista de verificaci3n PRISMA 2020 para resu´menes estructurados*

Secci3n/tema	Ítem n.º	Ítem de la lista de verificaci3n
TI TULO		
Título	1	Identifique el informe o publicaci3n como una revisi3n sistemática.
ANTECEDENTES		
Objetivos	2	Proporcione una declaraci3n explícita de los principales objetivos o preguntas que aborda la revisi3n.
ME´ TODOS		
Criterios de elegibilidad	3	Especifique los criterios de inclusi3n y exclusi3n de la revisi3n.
Fuentes de informaci3n	4	Especifique las fuentes de informaci3n (por ejemplo, bases de datos, registros) utilizadas para identificar los estudios y la fecha de la última b´queda en cada una de estas fuentes.
Riesgo de sesgo de los estudios individuales	5	Especifique los me´ todos utilizados para evaluar el riesgo de sesgo de los estudios individuales incluidos.
Síntesis de los resultados	6	Especifique los me´ todos utilizados para presentar y sintetizar los resultados.
RESULTADOS		
Estudios incluidos	7	Proporcione el número total de estudios incluidos y de participantes y resuma las características relevantes de los estudios.
Síntesis de los resultados	8	Presente los resultados de los desenlaces principales e indique, preferiblemente, el número de estudios incluidos y los participantes en cada uno de ellos. Si se ha realizado un metanálisis, indique el estimador de resumen y el intervalo de confianza o de credibilidad. Si se comparan grupos, describa la direcci3n del efecto (por ejemplo, que´ grupo se ha visto favorecido).
DISCUSIO´ N		
Limitaciones de la evidencia	9	Proporcione un breve resumen de las limitaciones de la evidencia incluida en la revisi3n (por ejemplo, riesgo de sesgo, inconsistencia –heterogeneidad– e imprecisi3n).
Interpretaci3n	10	Proporcione una interpretaci3n general de los resultados y sus implicaciones importantes.
OTROS		
Financiaci3n	11	Especifique la fuente principal de financiaci3n de la revisi3n.
Registro	12	Proporcione el nombre y el número de registro.

* Esta lista de verificaci3n conserva los mismos ítems que se incluyeron en la declaraci3n PRISMA para resu´menes publicada en 2013⁴⁸, pero ha sido revisada para que la redacci3n sea coherente con la declaraci3n PRISMA 2020. Adem3s, incluye un nuevo ítem que recomienda a los

autores que especifiquen los métodos utilizados para presentar y sintetizar los resultados (ítem n.8 6).

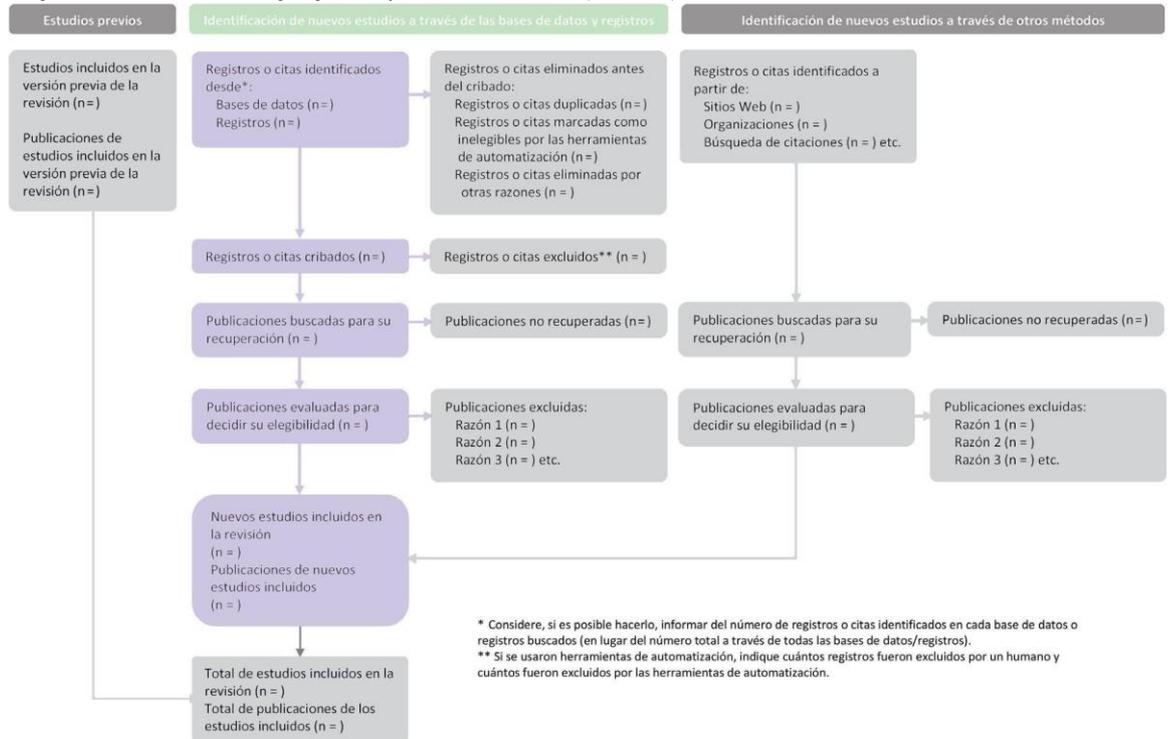


Figura 1. Diagrama de flujo PRISMA 2020. El nuevo diseño se ha adaptado a partir de los diagramas de flujo propuestos por Boers⁷⁰, Mayo-Wilson et al.⁷¹ y Stovold et al.⁷². Los recuadros en gris solo se deben completar si son aplicables; de lo contrario, deben eliminarse del diagrama de flujo. Observe que un «informe» puede ser un artículo de revista, una preimpresión, un resumen de conferencia, un registro de estudio, un informe de estudio clínico, una tesis/disertación, un manuscrito inédito, un informe gubernamental o cualquier otro documento que proporcione información

Anexo 3:Cosmin Risk of Bias check list

COSMIN Risk of Bias checklist

Date: December, 2017

Contact

L.B. Mokkink, PhD
VU University Medical Center
Department of Epidemiology and Biostatistics
Amsterdam Public Health research institute
P.O. box 7057
1007 MB Amsterdam
The Netherlands
Website: www.cosmin.nl
E-mail: w.mokkink@vumc.nl



How to cite the COSMIN Risk of Bias Checklist

Please refer to the following studies when using the COSMIN Risk of Bias Checklist:

Mokkink, L.B., De Vet, H.C.W., Prinsen, C.A.C, Patrick, D.L., Alonso, J., Bouter, L.M., et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Accepted for publication in Quality of Life Research.

Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., Vet, H. C., et al. COSMIN guideline for systematic reviews of Patient-Reported Outcome Measures. Submitted.

Terwee, C. B., Prinsen, C. A., Chiarotto, A., Vet, H. C., Westerman, M. J., Patrick, D. L., et al. COSMIN methodology for evaluating the content validity of Patient-Reported Outcome Measures: a Delphi study. Submitted.

For details on how to use the COSMIN risk of Bias checklist see 'COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) – user manual' and 'COSMIN methodology for assessing the content validity of Patient-Reported Outcome Measures (PROMs) - user manual' available from our website www.cosmin.nl.

Abbreviations used:

CTT – classical test theory

DIF – differential item functioning

IRT – Item response theory

MGCFAs – multi-group confirmatory factor analysis

MI – measurement invariance

NA – not applicable

PROM – patient-reported outcome measure

1PL model – 1 parameter IRT model

2PL model – 2 parameter IRT model

Instructions

Tick the boxes that need to be completed for the article

COSMIN Risk of Bias checklist	
	Box 1. PROM development
	Box 2. Content validity
	Box 3. Structural validity
	Box 4. Internal consistency
	Box 5. Cross-cultural validity\Measurement invariance
	Box 6. Reliability
	Box 7. Measurement error
	Box 8. Criterion validity
	Box 9. Hypotheses testing for construct validity
	Box 10. Responsiveness

To assess the methodological quality of each study, i.e. assessing the risk of bias of the result of a study, the corresponding COSMIN Risk of Bias box should be completed. To determine the overall quality of a study the lowest rating of any standard in the box is taken (i.e. “the worst score counts” principle). For example, if for a reliability study one item in a box is rated as ‘inadequate’, the overall methodological quality of that reliability study is rated as ‘inadequate’. The response option ‘NA’ (not applicable) is at issue for some standards. For example, when a study on structural validity is based on CTT, the standard on IRT is not applicable and this standard should not be considered in the “worst score counts”-rating for that specific study. For standards where this option is not at issue, these cells are grey and shouldn’t be used.

Box 1. PROM development						
1a. PROM design						
General design requirements		very good	adequate	doubtful	inadequate	NA
1	Is a clear description provided of the construct to be measured?	Construct clearly described			Construct not clearly described	
2	Is the origin of the construct clear: was a theory, conceptual framework or disease model used or clear rationale provided to define the construct to be measured?	Origin of the construct clear		Origin of the construct not clear		
3	Is a clear description provided of the target population for which the PROM was developed?	Target population clearly described			Target population not clearly described	
4	Is a clear description provided of the context of use	Context of use clearly described		Context of use not clearly described		
5	Was the PROM development study performed in a sample representing the target population for which the PROM was developed?	Study performed in a sample representing the target population	Assumable that the study was performed in a sample representing the target population, but not clearly described	Doubtful whether the study was performed in a sample representing the target population	Study not performed in a sample representing the target population (SKIP items 6-12)	

<i>Concept elicitation (relevance and comprehensiveness)</i>		very good	adequate	doubtful	inadequate	NA
6	Was an appropriate qualitative data collection method used to identify relevant items for a new PROM?	Widely recognized or well justified qualitative method used, suitable for the construct and study population	Assumable that the qualitative method was appropriate and suitable for the construct and study population, but not clearly described	Only quantitative (survey) method(s) used or doubtful whether the method was suitable for the construct and study population	Method used not appropriate or not suitable for the construct or study population	
7	Were skilled group moderators/interviewers used?	Skilled group moderators/ interviewers used	Group moderators /interviewers had limited experience or were trained specifically for the study	Not clear if group moderators /interviewers were trained or group moderators /interviewers not trained and no experience		Not applicable
8	Were the group meetings or interviews based on an appropriate topic or interview guide?	Appropriate topic or interview guide	Assumable that the topic or interview guide was appropriate, but not clearly described	Not clear if a topic guide was used or doubtful if topic or interview guide was appropriate or no guide		Not applicable
9	Were the group meetings or interviews recorded and transcribed verbatim?	All group meetings or interviews were recorded and transcribed verbatim	Assumable that all group meetings or interviews were recorded and transcribed verbatim, but not clearly described	Not clear if all group meetings or interviews were recorded and transcribed verbatim or recordings not transcribed verbatim or only notes were made during the group meetings/ interviews	No recording and no notes	Not applicable
10	Was an appropriate approach used to analyse the data?	A widely recognized or well justified approach was used	Assumable that the approach was appropriate, but not clearly described	Not clear what approach was used or doubtful whether the approach was appropriate	Approach not appropriate	
11	Was at least part of the data coded independently?	At least 50% of the data was coded by at least two researchers independently	11-49% of the data was coded by at least two researchers independently	Doubtful if two researchers were involved in the coding or only 1-10% of the data was coded by at least two researchers independently	Only one researcher was involved in coding or no coding	Not applicable
12	Was data collection continued until saturation was reached?	Evidence provided that saturation was reached	Assumable that saturation was reached	Doubtful whether saturation was reached	Evidence suggests that saturation was not reached	Not applicable
13	For quantitative studies (surveys): was the sample size appropriate?	≥100	50-99	30-49	<30	Not applicable

1b. <u>Cognitive interview study or other pilot test</u>		very good	adequate	doubtful	inadequate	NA
14	Was a cognitive interview study or other pilot test conducted? <i>General design requirements</i>	YES			NO (SKIP items 15-35)	
15	Was the cognitive interview study or other pilot test performed in a sample representing the target population? <i>Comprehensibility</i>	Study performed in a sample representing the target population	Assumable that the study was performed in a sample representing the target population, but not clearly described	Doubtful whether the study was performed in a sample representing the target population	Study not performed in a sample representing the target population	
16	Were patients asked about the <u>comprehensibility</u> of the PROM?	YES		NO (SKIP items 17-25)	Not clear (SKIP items 17-25)	
17	Were all items tested in their final form?	All items were tested in their final form	Assumable that all items were tested in their final form, but not clearly described	Not clear if all items were tested in their final form	Items were not tested in their final form or items were not re-tested after substantial adjustments	
18	Was an appropriate qualitative method used to assess the <u>comprehensibility</u> of the PROM instructions, items, response options, and recall period?	Widely recognized or well justified qualitative method used	Assumable that the method was appropriate but not clearly described	Only quantitative (survey) method(s) used or doubtful whether the method was appropriate or not clear if patients were asked about the comprehensibility of the items, response options or recall period or patients not asked about the comprehensibility of the PROM instructions or the recall period	Method used not appropriate or patients not asked about the comprehensibility of the items or the response options	
19	Was each item tested in an appropriate number of patients? For qualitative studies For quantitative (survey) studies	≥7 ≥50	4-6 ≥30	<4 or not clear <30 or not clear		
20	Were skilled interviewers used?	Skilled group moderators/ interviewers used	Group moderators /interviewers had limited experience or were trained specifically for the study	Not clear if group moderators /interviewers were trained or group moderators /interviewers not trained and no experience		Not applicable
21	Were the interviews based on an appropriate interview guide?	Appropriate topic or interview guide	Assumable that the topic or interview guide was appropriate, but not clearly described	Not clear if a topic guide was used or doubtful if topic or interview guide was appropriate or no guide		Not applicable

22	Were the interviews recorded and transcribed verbatim?	All group meetings or interviews were recorded and transcribed verbatim	Assumable that all group meetings or interviews were recorded and transcribed verbatim, but not clearly described	Not clear if all group meetings or interviews were recorded and transcribed verbatim or recordings not transcribed verbatim or only notes were made during the group meetings/ interviews	No recording and no notes	Not applicable
23	Was an appropriate approach used to analyse the data?	A widely recognized or well justified approach was used	Assumable that the approach was appropriate, but not clearly described	Not clear what approach was used or doubtful whether the approach was appropriate	Approach not appropriate	
24	Were at least two researchers involved in the analysis?	At least two researchers involved in the analysis	Assumable that at least two researchers were involved in the analysis, but not clearly described	Not clear if two researchers were included in the analysis or only one researcher involved in the analysis		
25	Were problems regarding the comprehensibility of the PROM instructions, items, response options, and recall period appropriately addressed by adapting the PROM?	No problems found or problems appropriately addressed and PROM was adapted and re-tested if necessary	Assumable that there were no problems or that problems were appropriately addressed, but not clearly described	Not clear if there were problems or doubtful if problems were appropriately addressed	Problems not appropriately addressed or PROM was adapted but items were not re-tested after substantial adjustments.	Not applicable

<i>Comprehensiveness</i>		very good	adequate	doubtful	inadequate	NA
26	Were patients asked about the <u>comprehensiveness</u> of the PROM?	YES		NO or not clear (SKIP items 27-35)		
27	Was the final set of items tested?	The final set of items was tested	Assumable that the final set of items was tested, but not clearly described	Not clear if the final set of items was tested or not the final set of items was tested or the set of items was not re-tested after items were removed or added		
28	Was an appropriate method used for assessing the <u>comprehensiveness</u> of the PROM?	Widely recognized or well justified method used	Assumable that the method was appropriate but not clearly described or only quantitative (survey) method(s) used	Doubtful whether the method was appropriate or method used not appropriate		
29	Was each item tested in an appropriate number of patients? For qualitative studies For quantitative (survey) studies	≥7 ≥50	4-6 ≥30	<4 or not clear <30 or not clear		
30	Were skilled interviewers used?	Skilled interviewers used	Interviewers had limited experience or were trained specifically for the study	Not clear if interviewers were trained or interviewers not trained and no experience		Not applicable

31	Were the interviews based on an appropriate interview guide?	Appropriate topic or interview guide	Assumable that the topic or interview guide was appropriate, but not clearly described	Not clear if a topic guide was used or doubtful if topic or interview guide was appropriate or no guide		Not applicable
32	Were the interviews recorded and transcribed verbatim?	All group meetings or interviews were recorded and transcribed verbatim	Assumable that all group meetings or interviews were recorded and transcribed verbatim, but not clearly described	Not clear if all group meetings or interviews were recorded and transcribed verbatim or recordings not transcribed verbatim or only notes were made during the group meetings/ interviews or no recording and no notes		Not applicable
33	Was an appropriate approach used to analyse the data?	A widely recognized or well justified approach was used	Assumable that the approach was appropriate, but not clearly described	Not clear what approach was used or doubtful whether the approach was appropriate or approach not appropriate		
34	Were at least two researchers involved in the analysis?	At least two researchers involved in the analysis	Assumable that at least two researchers were involved in the analysis, but not clearly described	Not clear if two researchers were included in the analysis or only one researcher involved in the analysis		
35	Were problems regarding the <u>comprehensiveness</u> of the PROM appropriately addressed by adapting the PROM?	No problems found or problems appropriately addressed and PROM was adapted and re-tested if necessary	Assumable that there were no problems or that problems were appropriately addressed, but not clearly described	Not clear if there were problems or doubtful if problems were appropriately addressed or PROM was adapted but items were not re-tested after substantial adjustments	Problems not appropriately addressed	Not applicable

Box 2. Content validity						
2a. Asking patients about relevance						
Design requirements						
	very good	adequate	doubtful	inadequate	NA	
1	Was an appropriate method used to ask patients whether each item is <u>relevant</u> for their experience with the condition?	Widely recognized or well justified method used	Only quantitative (survey) method(s) used or assumable that the method was appropriate but not clearly described	Not clear if patients were asked whether <u>each</u> item is relevant or doubtful whether the method was appropriate	Method used not appropriate or patients not asked about the relevance of all items	
2	Was each item tested in an appropriate number of patients? For qualitative studies For quantitative (survey) studies	≥7 ≥50	4-6 ≥30	<4 or not clear <30 or not clear		
3	Were skilled group moderators/interviewers used?	Skilled group moderators/interviewers used	Group moderators /interviewers had limited experience or were trained specifically for the study	Not clear if group moderators /interviewers were trained or group moderators /interviewers not trained and no experience		Not applicable
4	Were the group meetings or interviews based on an appropriate topic or interview guide?	Appropriate topic or interview guide	Assumable that the topic or interview guide was appropriate, but not clearly described	Not clear if a topic guide was used or doubtful if topic or interview guide was appropriate or no guide		Not applicable

5	Were the group meetings or interviews recorded and transcribed verbatim?	All group meetings or interviews were recorded and transcribed verbatim	Assumable that all group meetings or interviews were recorded and transcribed verbatim, but not clearly described	Not clear if all group meetings or interviews were recorded and transcribed verbatim or recordings not transcribed verbatim or only notes were made during the group meetings/ interviews	No recording and no notes	Not applicable
<i>Analyses</i>						
6	Was an appropriate approach used to analyse the data?	A widely recognized or well justified approach was used	Assumable that the approach was appropriate, but not clearly described	Not clear what approach was used or doubtful whether the approach was appropriate	Approach not appropriate	
7	Were at least two researchers involved in the analysis?	At least two researchers involved in the analysis	Assumable that at least two researchers were involved in the analysis, but not clearly described	Not clear if two researchers were included in the analysis or only one researcher involved in the analysis		

2b Asking patients about comprehensiveness						
<i>Design requirements</i>						
		very good	adequate	doubtful	inadequate	NA
8	Was an appropriate method used for assessing the <u>comprehensiveness</u> of the PROM?	Widely recognized or well justified method used	Only quantitative (survey) method(s) used or assumable that the method was appropriate but not clearly described	Doubtful whether the method was appropriate	Method used not appropriate	
9	Was each item tested in an appropriate number of patients? For qualitative studies For quantitative (survey) studies	≥7 ≥50	4-6 ≥30	<4 or not clear <30 or not clear		
10	Were skilled group moderators/interviewers used?	Skilled group moderators/ interviewers used	Group moderators /interviewers had limited experience or were trained specifically for the study	Not clear if group moderators /interviewers were trained or group moderators /interviewers not trained and no experience		Not applicable
11	Were the group meetings or interviews based on an appropriate topic or interview guide?	Appropriate topic or interview guide	Assumable that the topic or interview guide was appropriate, but not clearly described	Not clear if a topic guide was used or doubtful if topic or interview guide was appropriate or no guide		Not applicable

12	Were the group meetings or interviews recorded and transcribed verbatim?	All group meetings or interviews were recorded and transcribed verbatim	Assumable that all group meetings or interviews were recorded and transcribed verbatim, but not clearly described	Not clear if all group meetings or interviews were recorded and transcribed verbatim or recordings not transcribed verbatim or only notes were made during the group meetings/ interviews	No recording and no notes	Not applicable
<i>Analyses</i>						
13	Was an appropriate approach used to analyse the data?	A widely recognized or well justified approach was used	Assumable that the approach was appropriate, but not clearly described	Not clear what approach was used or doubtful whether the approach was appropriate	Approach not appropriate	
14	Were at least two researchers involved in the analysis?	At least two researchers involved in the analysis	Assumable that at least two researchers were involved in the analysis, but not clearly described	Not clear if two researchers were included in the analysis or only one researcher involved in the analysis		

2c Asking patients about comprehensibility						
Design requirements		very good	adequate	doubtful	inadequate	NA
15	Was an appropriate qualitative method used for assessing the <u>comprehensibility</u> of the PROM instructions, items, response options, and recall period?	Widely recognized or well justified qualitative method used	Assumable that the method was appropriate but not clearly described	Only quantitative (survey) method(s) used or doubtful whether the method was appropriate or not clear if patients were asked about the comprehensibility of the items, response options or recall period or patients not asked about the comprehensibility of the PROM instructions	Method used not appropriate or patients not asked about the comprehensibility of the items, response options, or recall period	
16	Was each item tested in an appropriate number of patients? For qualitative studies For quantitative (survey) studies	≥7 ≥50	4-6 ≥30	<4 or not clear <30 or not clear		
17	Were skilled group moderators/interviewers used?	Skilled group moderators/interviewers used	Group moderators/interviewers had limited experience or were trained specifically for the study	Not clear if group moderators/interviewers were trained or group moderators/interviewers not trained and no experience		
18	Were the group meetings or interviews based on an appropriate topic or interview guide?	Appropriate topic or interview guide	Assumable that the topic or interview guide was appropriate, but not clearly described	Not clear if a topic guide was used or doubtful if topic or interview guide was appropriate or no guide		Not applicable
19	Were the group meetings or interviews recorded and transcribed verbatim?	All group meetings or interviews were recorded and transcribed verbatim	Assumable that all group meetings or interviews were recorded and transcribed verbatim, but not clearly described	Not clear if all group meetings or interviews were recorded and transcribed verbatim or recordings not transcribed verbatim or only notes were made during the group meetings/interviews	No recording and no notes	Not applicable
Analyses						
20	Was an appropriate approach used to analyse the data?	A widely recognized or well justified approach was used	Assumable that the approach was appropriate, but not clearly described	Not clear what approach was used or doubtful whether the approach was appropriate	Approach not appropriate	
21	Were at least two researchers involved in the analysis?	At least two researchers involved in the analysis	Assumable that at least two researchers were involved in the analysis, but not clearly described	Not clear if two researchers were included in the analysis or only one researcher involved in the analysis		

2d. Asking professionals about relevance						
Design requirements		very good	adequate	doubtful	inadequate	NA
22	Was an appropriate method used to ask professionals whether each item is <u>relevant</u> for the construct of interest?	Widely recognized or well justified method used	Only quantitative (survey) method(s) used or assumable that the method was appropriate but not clearly described	Not clear if professionals were asked whether <u>each</u> item is relevant or doubtful whether the method was appropriate	Method used not appropriate or professionals not asked about the relevance of all items	
23	Were professionals from all relevant disciplines included?	Professionals from all required disciplines were included	Assumable that professionals from all required disciplines were included, but not clearly described	Doubtful whether professionals from all required disciplines were included or relevant professionals were not included		
24	Was each item tested in an appropriate number of professionals? For qualitative studies For quantitative (survey) studies	≥7 ≥50	4-6 ≥30	<4 or not clear <30 or not clear		
Analyses						
25	Was an appropriate approach used to analyse the data?	A widely recognized or well justified approach was used	Assumable that the approach was appropriate, but not clearly described	Not clear what approach was used or doubtful whether the approach was appropriate	Approach not appropriate	

26	Were at least two researchers involved in the analysis?	At least two researchers involved in the analysis	Assumable that at least two researchers were involved in the analysis, but not clearly described	Not clear if two researchers were included in the analysis or only one researcher involved in the analysis		
----	---	---	--	--	--	--

2e. Asking professionals about comprehensiveness						
Design requirement		very good	adequate	doubtful	inadequate	NA
27	Was an appropriate method used for assessing the <u>comprehensiveness</u> of the PROM?	Widely recognized or well justified method used	Only quantitative (survey) method(s) used or assumable that the method was appropriate but not clearly described	Doubtful whether the method was appropriate	Method used not appropriate	
28	Were professionals from all relevant disciplines included?	Professionals from all required disciplines were included	Assumable that professionals from all required disciplines were included, but not clearly described	Doubtful whether professionals from all required disciplines were included or relevant professionals were not included		
29	Was each item tested in an appropriate number of professionals? For qualitative studies For quantitative (survey) studies	≥7 ≥50	4-6 ≥30	<4 or not clear <30 or not clear		
Analyses						
30	Was an appropriate approach used to analyse the data?	A widely recognized or well justified approach was used	Assumable that the approach was appropriate, but not clearly described	Not clear what approach was used or doubtful whether the approach was appropriate	Approach not appropriate	

31	Were at least two researchers involved in the analysis?	At least two researchers involved in the analysis	Assumable that at least two researchers were involved in the analysis, but not clearly described	Not clear if two researchers were included in the analysis or only one researcher involved in the analysis	
----	---	---	--	--	--

Box 3. Structural validity						
Does the scale consist of effect indicators, i.e. is it based on a reflective model? ¹ yes / no						
Does the study concern unidimensionality or structural validity? ² unidimensionality / structural validity						
<i>Statistical methods</i>		very good	adequate	doubtful	inadequate	NA
1	For CTT: Was exploratory or confirmatory factor analysis performed?	Confirmatory factor analysis performed	Exploratory factor analysis performed		No exploratory or confirmatory factor analysis performed	Not applicable
2	For IRT/Rasch: does the chosen model fit to the research question?	Chosen model fits well to the research question	Assumable that the chosen model fits well to the research question	Doubtful if the chosen model fits well to the research question	Chosen model does not fit to the research question	Not applicable
3	Was the sample size included in the analysis adequate?	FA: 7 times the number of items and ≥ 100 Rasch/1PL models: ≥ 200 subjects 2PL parametric IRT models OR Mokken scale analysis: ≥ 1000 subjects	FA: at least 5 times the number of items and ≥ 100 ; OR at least 6 times number of items but < 100 Rasch/1PL models: 100-199 subjects 2PL parametric IRT models OR Mokken scale analysis: 500-999 subjects	FA: 5 times the number of items but < 100 Rasch/1PL models: 50-99 subjects 2PL parametric IRT models OR Mokken scale analysis: 250-499 subjects	FA: < 5 times the number of items Rasch/1PL models: < 50 subjects 2PL parametric IRT models OR Mokken scale analysis: < 250 subjects	
<i>Other</i>						
4	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws (e.g. rotation method not described)	Other important methodological flaws (e.g. inappropriate rotation method)	

¹ If the scale is not based on a reflective model, unidimensionality or structural validity is not relevant.

² In a systematic review, it is helpful to make a distinction between studies where factor analysis is performed on each (sub)scale separately to evaluate whether the (sub)scales are unidimensional (unidimensionality studies) and studies where factor analysis is performed on all items of an instrument to evaluate the (expected) number of subscales in the instrument and the clustering of items within subscales (structural validity studies).

Box 4. Internal consistency						
Does the scale consist of effect indicators, i.e. is it based on a reflective model? ¹ yes / no						
<i>Design requirements</i>		very good	adequate	doubtful	inadequate	NA
1 Was an internal consistency statistic calculated for each unidimensional scale or subscale separately?		Internal consistency statistic calculated for each unidimensional scale or subscale		Unclear whether scale or sub scale is unidimensional	Internal consistency statistic NOT calculated for each unidimensional scale or sub scale	
<i>Statistical methods</i>						
2 For continuous scores: Was Cronbach's alpha or omega calculated?		Cronbach's alpha, or Omega calculated		Only item-total correlations calculated	No Cronbach's alpha and no item-total correlations calculated	Not applicable
3 For dichotomous scores: Was Cronbach's alpha or KR-20 calculated?		Cronbach's alpha or KR-20 calculated		Only item-total correlations calculated	No Cronbach's alpha or KR-20 and no item-total correlations calculated	Not applicable
4 For IRT-based scores: Was standard error of the theta (SE(θ)) or reliability coefficient of estimated latent trait value (index of (subject or item) separation) calculated?		SE(θ) or reliability coefficient calculated			SE(θ) or reliability coefficient NOT calculated	Not applicable
<i>Other</i>						
5 Were there any other important flaws in the design or statistical methods of the study?		No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

¹ If the scale is not based on a reflective model, internal consistency is not relevant

Box 5. Cross-cultural validity\Measurement invariance						
<i>Design requirements</i>		very good	adequate	doubtful	inadequate	NA
1 Were the samples similar for relevant characteristics except for the group variable?		Evidence provided that samples were similar for relevant characteristics except group variable	Stated (but no evidence provided) that samples were similar for relevant characteristics except group variable	Unclear whether samples were similar for relevant characteristics except group variable	Samples were NOT similar for relevant characteristics except group variable	
<i>Statistical methods</i>						
2 Was an appropriate approach used to analyse the data?		A widely recognized or well justified approach was used	Assumable that the approach was appropriate, but not clearly described	Not clear what approach was used or doubtful whether the approach was appropriate	Approach not appropriate	Not applicable
3 Was the sample size included in the analysis adequate?		Regression analyses or IRT/Rasch based analyses: 200 subjects per group	150 subjects per group	100 subjects per group	< 100 subjects per group	
		MGCFA*: 7 times the number of items and ≥ 100	5 times the number of items and ≥ 100 ; OR 5-7 times the number of items but <100	5 times the number of items but <100	<5 times the number of items	
<i>Other</i>						
4 Were there any other important flaws in the design or statistical methods of the study?		No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

*MGCFA: multi-group confirmatory factor analyses

Box 6. Reliability		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Evidence provided that patients were stable	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?	ICC calculated and model or formula of the ICC is described	ICC calculated but model or formula of the ICC not described or not optimal. Pearson or Spearman correlation coefficient calculated with evidence provided that no systematic change has occurred	Pearson or Spearman correlation coefficient calculated WITHOUT evidence provided that no systematic change has occurred or WITH evidence that systematic change has occurred	No ICC or Pearson or Spearman correlations calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was kappa calculated?	Kappa calculated			No kappa calculated	Not applicable
6	For ordinal scores: Was a weighted kappa calculated?	Weighted Kappa calculated		Unweighted Kappa calculated or not described		Not applicable
7	For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic	Weighting scheme described	Weighting scheme NOT described			Not applicable
<i>Other</i>						
8	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 7. Measurement error		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
1	Were patients stable in the interim period on the construct to be measured?	Patients were stable (evidence provided)	Assumable that patients were stable	Unclear if patients were stable	Patients were NOT stable	
2	Was the time interval appropriate?	Time interval appropriate		Doubtful whether time interval was appropriate or time interval was not stated	Time interval NOT appropriate	
3	Were the test conditions similar for the measurements? (e.g. type of administration, environment, instructions)	Test conditions were similar (evidence provided)	Assumable that test conditions were similar	Unclear if test conditions were similar	Test conditions were NOT similar	
<i>Statistical methods</i>						
4	For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?	SEM, SDC, or LoA calculated	Possible to calculate LoA from the data presented		SEM calculated based on Cronbach's alpha, or on SD from another population % agreement not calculated	Not applicable
5	For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated?	% positive and negative agreement calculated	% agreement calculated			Not applicable
<i>Other</i>						
6	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 8. Criterion validity		very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>						
1	For continuous scores: Were correlations, or the area under the receiver operating curve calculated?	Correlations or AUC calculated			Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scores: Were sensitivity and specificity determined?	Sensitivity and specificity calculated			Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>						
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Box 9. Hypotheses testing for construct validity		very good	adequate	doubtful	inadequate	NA
9a. Comparison with other outcome measurement instruments (convergent validity)						
<i>Design requirements</i>						
1	Is it clear what the comparator instrument(s) measure(s)?	Constructs measured by the comparator instrument(s) is clear			Constructs measured by the comparator instrument(s) is not clear	
2	Were the measurement properties of the comparator instrument(s) sufficient?	Sufficient measurement properties of the comparator instrument(s) in a population similar to the study population	Sufficient measurement properties of the comparator instrument(s) but not sure if these apply to the study population	Some information on measurement properties of the comparator instrument(s) in any study population	No information on the measurement properties of the comparator instrument(s), OR evidence for insufficient measurement properties of the comparator instrument(s)	
<i>Statistical methods</i>						
3	Was the statistical method appropriate for the hypotheses to be tested?	Statistical method was appropriate	Assumable that statistical method was appropriate	Statistical method applied NOT optimal	Statistical method applied NOT appropriate	

4		Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws (e.g. only data presented on a comparison with an instrument that measures another construct)	Other important methodological flaws	
---	--	---	---	--	--	--------------------------------------	--

9b. Comparison between subgroups (discriminative or known-groups validity)		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
5	Was an adequate description provided of important characteristics of the subgroups?	Adequate description of the important characteristics of the subgroups	Adequate description of most of the important characteristics of the subgroups	Poor or no description of the important characteristics of the subgroups		
<i>Statistical methods</i>						
6	Was the statistical method appropriate for the hypotheses to be tested?	Statistical method was appropriate	Assumable that statistical method was appropriate	Statistical method applied NOT optimal	Statistical method applied NOT appropriate	
<i>Other</i>						
7	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws (e.g. only data presented on a comparison with an instrument that measures another construct)	Other important methodological flaws	

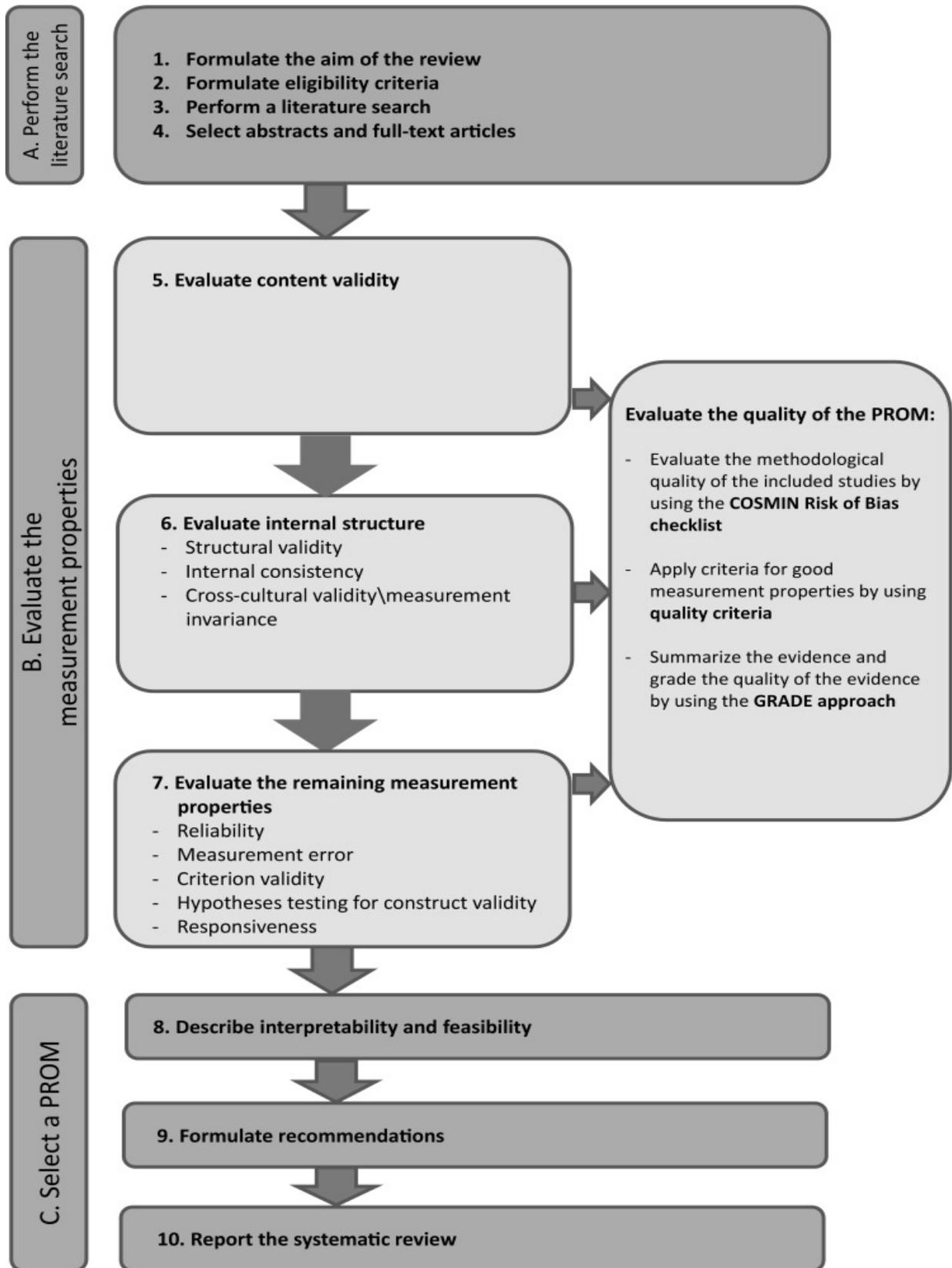
Box 10. Responsiveness					
10a. Criterion approach (i.e. comparison to a gold standard)					
	very good	adequate	doubtful	inadequate	NA
<i>Statistical methods</i>					
1	For continuous scores: Were correlations between change scores, or the area under the Receiver Operator Curve (ROC) curve calculated?	Correlations or Area under the ROC Curve (AUC) calculated		Correlations or AUC NOT calculated	Not applicable
2	For dichotomous scales: Were sensitivity and specificity (changed versus not changed) determined?	Sensitivity and specificity calculated		Sensitivity and specificity NOT calculated	Not applicable
<i>Other</i>					
3	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws	Other minor methodological flaws	Other important methodological flaws	

10b. Construct approach (i.e. hypotheses testing; comparison with other outcome measurement instruments)					
	very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>					
4	Is it clear what the comparator instrument(s) measure(s)?	Constructs measured by the comparator instrument(s) is clear		Constructs measured by the comparator instrument(s) is not clear	
5	Were the measurement properties of the comparator instrument(s) sufficient?	Sufficient measurement properties of the comparator instrument(s) in a population similar to the study population	Sufficient measurement properties of the comparator instrument(s) but not sure if these apply to the study population	Some information on measurement properties of the comparator instrument(s) in any study population	NO information on the measurement properties of the comparator instrument(s) OR evidence of poor quality of comparator instrument(s)
<i>Statistical methods</i>					
6	Was the statistical method appropriate for the hypotheses to be tested?	Statistical method was appropriate	Assumable that statistical method were appropriate	Statistical method applied NOT optimal	Statistical method applied NOT appropriate
<i>Other</i>					
7	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws	Other minor methodological flaws	Other important methodological flaws	

10c. Construct approach: (i.e. hypotheses testing: comparison between subgroups)						
		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
8	Was an adequate description provided of important characteristics of the subgroups?	Adequate description of the important characteristics of the subgroups	Adequate description of most of the important characteristics of the subgroups	Poor or no description of the important characteristics of the subgroups		
<i>Statistical methods</i>						
9	Was the statistical method appropriate for the hypotheses to be tested?	Statistical method was appropriate	Assumable that statistical method was appropriate	Statistical method applied NOT optimal	Statistical method applied NOT appropriate	
<i>Other</i>						
10	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

10d. Construct approach: (i.e. hypotheses testing: before and after intervention)						
		very good	adequate	doubtful	inadequate	NA
<i>Design requirements</i>						
11	Was an adequate description provided of the intervention given?	Adequate description of the intervention		Poor description of the intervention	NO description of the intervention	
<i>Statistical methods</i>						
12	Was the statistical method appropriate for the hypotheses to be tested?	Statistical method was appropriate	Assumable that statistical method was appropriate	Statistical method applied NOT optimal	Statistical method applied NOT appropriate	
<i>Other</i>						
13	Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Anexo 4. COSMIN guideline for Systematic reviews of PROMS



Anexo 5.Evaluación calidad de estudio:

Spanish Version of the Avoidance-Endurance Questionnaire: Factor Structure and Psychometric Properties. Ruiz-Párraga et al.(2015)

Criterio GRADE	Calificación (Marcar 1)	Notas explicando las razones de la disminución o aumento de calidad	Calidad de la evidencia
Diseño de estudio	RCT o NO RCT	Diseño observacional transversal	Moderada
Riesgo de Sesgo	No hay Serio(-1) Muy serio(-2)	La mayoría de los parámetros son adecuados o no se mencionan.	
Inconsistencia	No hay Serio(-1) Muy serio(-2)	No hay más evidencia disponible para contrastar la consistencia del contenido.	
Indirectez	No hay Serio(-1) Muy serio(-2)	El cuestionario se puede referir al contexto que se plantea dentro del artículo.	
Imprecisión	No hay Serio(-1) Muy serio(-2)	La confiabilidad interna es aceptable y se realiza sobre 150 personas, es estadísticamente significativo.	
Sesgo de publicación	Indetectable Fuertemente sospechoso	No se analiza realmente la selección aparentemente son personas aleatorias que refieren dolor crónico	
Otros	Gran efecto(+1 o +2) Sin factores de confusión(+1 o +2) Responde a lo esperado(+1 o+2)	Aparentemente el artículo evalúa la herramienta como confiable en el lenguaje hispano	

Juicio clínico: La calidad de la evidencia es moderada, esto indica que el cuestionario mide lo que se espera que mida. El hecho de que haya tan poca evidencia al respecto sobre él en esta lengua, requiere de más investigación para confirmar su juicio. Además el AEQ-SP muestra Correlaciones moderadas en variables de dolor y discapacidad lo que puede que sea sensible al cambio de resultados. Se debe considerar esta como herramienta de evaluación integral dentro del tratamiento de pacientes con dolor crónico de cuello. Los datos de confiabilidad interna superan el $r=0,83$, en cuanto a validez de constructo y criterio los datos fueron bajos.

Recomendaciones:

-Se recomienda seguir abriendo vías de investigación para seguir evaluando el AEQ-SP.

-El AEQ-SP se puede usar como parte de una evaluación integral para evaluar la evitación y la resistencia al dolor en pacientes con dolor crónico.

-El AEQ-SP puede ser una herramienta útil para los pacientes que desean comprender mejor su propia evitación y resistencia al dolor. Deben usar de un profesional para su interpretación y comprensión.

Psychometric characteristics of the Spanish version of instruments to measure neck pain disability. M Kovacs ,et al.(2008)

Criterio GRADE	Calificación (Marcar 1)	Notas explicando las razones de la disminución o aumento de calidad	Calidad de la evidencia
Diseño de estudio	RCT o NO RCT	Diseño observacional transversal	Baja
Riesgo de Sesgo	No hay Serio(-1) Muy serio(-2)	El estudio no menciona aleatorización ni cegamiento, incrementando el riesgo de sesgo en selección y medición.	
Inconsistencia	No hay Serio(-1) Muy serio(-2)	No se describe variabilidad entre estudios, limitando la capacidad para confirmar hallazgos.	
Indirectez	No hay Serio(-1) Muy serio(-2)	El cuestionario se puede referir al contexto que se plantea dentro del artículo.	
Imprecisión	No hay Serio(-1) Muy serio(-2)	La confiabilidad interna es adecuada y se realiza sobre 221 personas, es estadísticamente significativo, aunque aparecen algunas comparaciones sospechosas.	
Sesgo de publicación	Indetectable Fuertemente sospechoso	Se realiza una selección de pacientes con sintomatología en diferentes regiones y de forma sucesiva.	
Otros	Gran efecto(+1 o +2) Sin factores de confusión(+1 o +2) Responde a lo esperado(+1 o+2)	Aparentemente el artículo evalúa la herramienta como confiable en el lenguaje hispano. La diferencia entre el NDI y el NPQ es considerable, lo que fortalece la evidencia.	

Juicio clínico: La baja calidad de la evidencia limita la confianza en los resultados del estudio. El NDI presenta una ventaja sobre el COM en términos de efectividad y precisión. El COM puede sobreestimar la mejoría del paciente, por lo que es precoz formular cambios en la práctica clínica basándose en esta información.

Recomendaciones:

-Las herramientas NDI y NPQ salvo la baja calidad de la evidencia muestran mejores valores en la evaluación y uso sobre personas con patología de cuello.

-Sería conveniente realizar un ensayo controlado aleatorio o estudios de cohortes prospectivos, para comparar directamente el NDI y el COM en la evaluación de la discapacidad relacionada con el dolor de cuello

Craniofacial pain and disability inventory (CF-PDI): development and psychometric validation of a new questionnaire. La Touche, et al. (2014)

Criterio GRADE	Calificación (Marcar 1)	Notas explicando las razones de la disminución o aumento de calidad	Calidad de la evidencia
Diseño de estudio	RCT o NO RCT	Diseño de estudio observacional prospectivo, transversal y descriptivo de validación.	Moderada
Riesgo de Sesgo	No hay Serio(-1) Muy serio(-2)	El estudio menciona una selección sistemática sobre una población específica conveniente que puede no representar a la población con sintomatología general	
Inconsistencia	No hay Serio(-1) Muy serio(-2)	Los índices de cronbach y la confiabilidad del test-retest justifican la calificación.	
Indirectez	No hay Serio(-1) Muy serio(-2)	El cuestionario se dirige específicamente a personas con dolor crónico craneofacial. Analiza sus datos psicométricos y comprensibilidad.	
Imprecisión	No hay Serio(-1) Muy serio(-2)	La confiabilidad interna es adecuada y se realiza sobre 192 personas, aunque puede aparecer sospechoso en algunos subanálisis.	
Sesgo de publicación	Indetectable Fuertemente sospechoso	No perceptibles si se excluyen datos.	
Otros	Gran efecto(+1 o +2) Sin factores de confusión(+1 o +2) Responde a lo esperado(+1 o +2)	El estudio presenta detalles estadísticos con precisión, incluyendo medias, desviaciones estándar, y coeficientes de variación. Aplicó técnicas estadísticas avanzadas para validar el instrumento, incluyendo análisis factorial y análisis de consistencia interna.	

Juicio clínico: Las propiedades psicométricas del CF-PDI parecen prometedoras según la consistencia interna, la confiabilidad de la prueba-retest y la validez de constructo informadas. Sería conveniente diseñar un estudio randomizado aleatorio que respondiera a este en consecuencia a efectos de tratamiento o como medida del mismo.

Recomendaciones:

- El CF-PDI puede ser una herramienta útil para medir el dolor, la discapacidad y el estado funcional en pacientes con dolor craneofacial.
- Se necesita más investigación utilizando diseños de estudios más sólidos (por ejemplo, ECA) para confirmar su efectividad.

The Catalan version of the Pain Catastrophizing Scale: a useful instrument to assess catastrophic thinking in whiplash patients. Miró et al. (2007)

Criterio GRADE	Calificación (Marcar 1)	Notas explicando las razones de la disminución o aumento de calidad	Calidad de la evidencia
Diseño de estudio	RCT o NO RCT	Estudio Observacional, transversal y de cohorte prospectiva	Moderada
Riesgo de Sesgo	No hay Serio(-1) Muy serio(-2)	La muestra es seleccionada por conveniencia y se excluye población que podría ser representativa. Hay riesgo de subjetividad en la medición de los valores.	
Inconsistencia	No hay Serio(-1) Muy serio(-2)	Se ha evaluado y tiene niveles de consistencia adecuados	
Indirectez	No hay Serio(-1) Muy serio(-2)	El estudio muestra dos direcciones de evaluación estudiante universitarios y personas con latigazo cervical. Determinando la población diana de estudio.	
Imprecisión	No hay Serio(-1) Muy serio(-2)	El estudio proporciona información detallada sobre los análisis estadísticos realizados.	
Sesgo de publicación	Indetectable Fuertemente sospechoso	Cierto que se selecciona una población específica, no se determina mucho más los criterios de selección. Al ser una gran muestra se determina que se aleatoriza la selección.	
Otros	Gran efecto(+1 o +2) Sin factores de confusión(+1 o +2) Responde a lo esperado(+1 o +2)	No rasgos considerables	

Juicio clínico: La PCS-Cat muestra fuerza para evaluar el catastrofismo del dolor, pero se recomienda precaución debido a las limitaciones de la evidencia. Se necesitan más estudios para investigar para mejorar su consistencia interna de algunas subescalas y replicar los resultados en muestras más grandes de pacientes con latigazo cervical.

Recomendaciones: La PCS-Cat puede ser una herramienta útil para evaluar el catastrofismo del dolor, pero se recomienda su uso con precaución y en combinación con otras herramientas de evaluación clínica. El catastrofismo del dolor es un factor importante que puede influir en la recuperación de pacientes con dolor crónico, incluyendo el latigazo cervical.

Validation of a Spanish version of the Spine Functional Index Cuesta-Vargas et al.(2014)

Criterio GRADE	Calificación (Marcar 1)	Notas explicando las razones de la disminución o aumento de calidad	Calidad de la evidencia
Diseño de estudio	RCT o NO RCT	Estudio Observacional prospectivo	Moderada
Riesgo de Sesgo	No hay Serio(-1) Muy serio(-2)	No se observa un riesgo de sesgo significativo en el estudio, ya que se empleó un diseño observacional de dos etapas con medidas para minimizar el sesgo, como la asignación aleatoria de participantes a grupos de estudio y la ceguera de los evaluadores a los puntajes iniciales.	
Inconsistencia	No hay Serio(-1) Muy serio(-2)	Se utilizaron métodos estándar para la traducción y adaptación cultural del cuestionario, y se llevaron a cabo análisis estadísticos robustos para evaluar las propiedades psicométricas del instrumento.	
Indirectez	No hay Serio(-1) Muy serio(-2)	Se realizaron esfuerzos para garantizar que la traducción y adaptación cultural del cuestionario fueran apropiadas para la población objetivo	
Imprecisión	No hay Serio(-1) Muy serio(-2)	Se emplearon métodos estadísticos adecuados y se calcularon intervalos de confianza para estimar la precisión de las medidas.	
Sesgo de publicación	Indetectable Fuertemente sospechoso	Se proporciona una descripción detallada de los métodos utilizados y los resultados obtenidos, lo que sugiere que se incluyeron y reportaron todos los resultados relevantes.	
Otros	Gran efecto(+1 o +2) Sin factores de confusión(+1 o +2) Responde a lo esperado(+1 o+2)	Ya que el estudio proporciona evidencia sólida sobre las propiedades psicométricas del SFI-Sp en la población de habla hispana.El estudio establece criterios claros de inclusión y exclusión, utiliza métodos estandarizados para la traducción y validación del cuestionario.El estudio encuentra que el SFI-Sp tiene una alta consistencia interna, una buena fiabilidad y validez de criterio, lo que es consistente con las expectativas teóricas y con los hallazgos de estudios previos sobre el SFI en otras poblaciones.	

Juicio clínico: El SFI-Sp puede ayudar a medir la función de la columna vertebral y el progreso del tratamiento en pacientes con dolor lumbar no específico. Es una herramienta validada y confiable para la población de habla hispana. Puede ser útil para el seguimiento de los pacientes a lo largo del tiempo y para evaluar la efectividad de las intervenciones. La herramienta es fácil de administrar e interpretar además no requiere entrenamiento especializado para su uso y puede ser completado por los pacientes de forma rápida y sencilla.

Recomendaciones: Se debe considerar utilizar el Spine Functional Index traducido al español (SFI-Sp) como una medida confiable y válida de la función espinal en pacientes de habla hispana con dolor de espalda o cuello. Se sugiere realizar más investigaciones para validar aún más el SFI-Sp en diferentes contextos clínicos y poblaciones, así como para evaluar su sensibilidad al cambio.

Validation of a Spanish version of the Neck Disability Index Ortega et al.(2007)

Criterio GRADE	Calificación (Marcar 1)	Notas explicando las razones de la disminución o aumento de calidad	Calidad de la evidencia
Diseño de estudio	RCT o NO RCT	El diseño de estudio es Observacional y prospectivo	Moderado
Riesgo de Sesgo	No hay Serio(-1) Muy serio(-2)	Los pacientes fueron reclutados en un departamento de rehabilitación y en un departamento de prevención de riesgos laborales. Esto podría introducir sesgo de selección	
Inconsistencia	No hay Serio(-1) Muy serio(-2)	No se informa de diferencia entre los subgrupos	
Indirectez	No hay Serio(-1) Muy serio(-2)	No se acaba de percibir si el dolor aparecido proviene del ámbito laboral o general.	
Imprecisión	No hay Serio(-1) Muy serio(-2)	La muestra de estudio es muy baja y poco confiable.	
Sesgo de publicación	Indetectable Fuertemente sospechoso	Se proporciona una descripción detallada de los métodos utilizados y los resultados obtenidos	
Otros	Gran efecto(+1 o +2) Sin factores de confusión(+1 o +2) Responde a lo esperado(+1 o+2)	El artículo es claro, determina bien la metodología y como se reevaluará la misma y los índices de los resultados.	

Juicio Clínico:

El estudio demuestra que la versión española del NDI es una herramienta confiable y válida para evaluar la discapacidad cervical en pacientes con dolor cervical. La aplicabilidad de los resultados a la población general con dolor cervical podría limitarse debido al sesgo de selección en el reclutamiento de pacientes. Se necesitan estudios adicionales para evaluar la validez y confiabilidad del NDI y hacer evaluaciones cruzadas.

Recomendaciones:

-La versión española del NDI puede ser una herramienta útil para evaluar la discapacidad cervical en pacientes con dolor cervical.

-Se necesitan estudios adicionales para evaluar la validez y confiabilidad del NDI en otras poblaciones

Spanish version of the screening Örebro musculoskeletal pain questionnaire: a cross-cultural adaptation and validation Cuesta-Vargas et al. (2014)

Criterio GRADE	Calificación (Marcar 1)	Notas explicando las razones de la disminución o aumento de calidad	Calidad de la evidencia
Diseño de estudio	RCT o NO RCT	El diseño de estudio es Observacional y prospectivo	Moderado-Alto
Riesgo de Sesgo	No hay Serio(-1) Muy serio(-2)	Los criterios de selección parecen estar bien definidos.	
Inconsistencia	No hay Serio(-1) Muy serio(-2)	Los valores de consistencia son adecuados.	
Indirectez	No hay Serio(-1) Muy serio(-2)	La población y la herramienta están correctamente interrelacionadas.	
Imprecisión	No hay Serio(-1) Muy serio(-2)	El número de pacientes es adecuado, pero existe variación entre población diana y objetivo de la PROM	
Sesgo de publicación	Indetectable Fuertemente sospechoso	El sesgo de publicación es indetectable porque el estudio parece haber sido conducido de manera transparente y completa.	
Otros	Gran efecto(+1 o +2) Sin factores de confusión(+1 o +2) Responde a lo esperado(+1 o+2)	El estudio muestra un gran efecto, no hay factores de confusión y los resultados son coherentes con lo esperado, lo que aumenta la confianza en la validez de los hallazgos.	

Juicio clínico: La versión española del OMPQ muestra potencial como herramienta para identificar pacientes con dolor musculoesquelético en riesgo de desarrollar discapacidad crónica. Sin embargo, las limitaciones del estudio, especialmente en cuanto a la generalización de los resultados y la falta de validez predictiva, requieren investigaciones adicionales para confirmar su utilidad clínica.

Recomendaciones:

-Se necesitan estudios de replicación con muestras más grandes y diversas para confirmar la estructura factorial del OMPQ.

-Se recomienda investigar la dimensión "Afrontamiento" para mejorar su confiabilidad.

-Es fundamental realizar estudios de validez predictiva para evaluar la capacidad del OMPQ de predecir la discapacidad crónica en la población española.

-Se sugiere analizar los resultados por separado para pacientes con dolor agudo y crónico.

Confirmatory factor analysis of the neck disability index, comparing patients with whiplash associated disorders to a control group with non-specific neck pain. Gabel et al. (2014)

Criterio GRADE	Calificación (Marcar 1)	Notas explicando las razones de la disminución o aumento de calidad	Calidad de la evidencia
Diseño de estudio	RCT o NO RCT	Estudio Observacional	Moderado
Riesgo de Sesgo	No hay Serio(-1) Muy serio(-2)	El artículo menciona limitaciones en la recolección de datos original, pero no evalúa el sesgo en el análisis secundario.	
Inconsistencia	No hay Serio(-1) Muy serio(-2)	Los resultados del modelo unidimensional del NDI son consistentes en todos los subgrupos analizados.	
Indirectez	No hay Serio(-1) Muy serio(-2)	El estudio se centra en una validez de constructo y no criterio eso da dudas sobre los posibles resultados.	
Imprecisión	No hay Serio(-1) Muy serio(-2)	El tamaño de la muestra combinada (n = 1755) es muy buena para proporcionar estimaciones precisas	
Sesgo de publicación	Indetectable Fuertemente sospechoso	No se menciona si se abordó el sesgo de publicación en el análisis.	
Otros	Gran efecto(+1 o +2) Sin factores de confusión(+1 o +2) Responde a lo esperado(+1 o+2)	El NDI tiene un gran efecto y el estudio no genera confusiones a gran escala.	

Juicio clínico:

Este análisis secundario sugiere que el NDI tiene una estructura factorial unidimensional consistente en pacientes con dolor cervical, tanto WAD como no-WAD, y en ambos sexos. Sin embargo, la naturaleza observacional del estudio y su enfoque en la estructura interna del NDI reducen la fuerza de la evidencia. Los clínicos pueden considerar el uso del NDI, pero deben tener en cuenta sus limitaciones y considerar otras medidas de discapacidad cervical que aborden directamente la validez de criterio.

Recomendaciones:

-El estudio proporciona evidencia a favor de la validez del NDI como medida unidimensional de la discapacidad cervic

Anexo 6. Evaluación cuestionario AEQ(artículo de Hasenbring):

Table 3

Final factor solution with factor loadings for the cognitive items, mean and standard deviation (SD) for each item.

Description of item	KPI-CRPS sub-scale	Factor I Help/- hopelessness	Factor II Thought suppression	Factor III Catastrophizing	Mean	SD
1. Why do I have to bear this heavy burden?	HHS	0.63	0.14	0.18	1.71	1.18
2. It seems the pain will never ease up.	HHS	0.82	0.06	-0.15	2.61	1.18
3. This damn pain spoils everything.	HHS	0.84	-0.05	-0.07	2.36	2.18
6. I will not be able to endure it any longer.	HHS	0.77	0.03	-0.01	1.88	1.18
8. Oh, it is not going to get any better.	HHS	0.89	-0.03	-0.09	2.39	1.18
9. Nothing helps anymore!	HHS	0.77	-0.06	-0.10	1.55	1.18
11. Life is hardly worth living with pain like this.	HHS	0.68	-0.01	0.11	0.97	1.18
12. Whatever will I do if it gets worse again?	HHS	0.66	-0.11	0.15	2.33	1.18
14. How long do I have to put up with pain like this?	HHS	0.75	0.05	0.16	2.17	1.18
5. It is important for me now to hold on!	TSS	0.06	0.74	0.04	3.49	1.18
10. Pull yourself together!	TSS	0.05	0.84	-0.15	3.46	1.18
13. Don't make such a fuss!	TSS	-0.25	0.81	0.05	2.56	1.18
16. It is important not to let myself go now.	TSS	0.13	0.76	0.06	2.84	1.18
4. I can't have a tumor, can I?	CTS	0.00	0.04	0.77	0.58	1.18
7. I wonder if I have the same serious illness as..	CTS	-0.03	-0.05	0.84	2.33	1.18
15. It is not a serious illness, is it?	CTS	-0.01	0.00	0.89	2.17	1.18
Explained variance in percent		37.63%	14.56%	10.72%		

Table 4

Final factor solution with factor loadings for the behavioral items, mean and standard deviation (SD) for each item

Item/description of item	KPI-CoPRS sub-scale	Factor I Avoidance of social activities	Factor II Avoidance of physical activities	Factor III Humour and distraction	Factor IV Pain persistence	Mean	SD
2. I avoid visiting my friends.	ASAS	0.76	0.12	-0.05	0.14	2.73	1.18
7. I cancel private appointments.	ASAS	0.83	0.08	-0.08	0.16	2.78	1.18
8. I cancel a visit to an event.	ASAS	0.89	0.06	0.29	-0.06	1.96	1.18
14. I break off meeting with friends.	ASAS	0.94	-0.07	0.03	0.03	1.92	1.18
18. I call guests to cancel invitations.	ASAS	0.91	0.01	0.11	-0.01	1.64	1.18
21. I avoid other people's company.	ASAS	0.82	-0.08	-0.08	0.15	2.31	1.18
1. I stop physical demanding activities.	APAS	-0.08	0.88	0.12	0.03	4.26	1.18
3. I take a rest.	APAS	0.09	0.70	0.03	-0.12	3.94	1.18
9. I avoid physical strenuous activities.	APAS	-0.06	0.86	-0.08	0.05	3.99	1.18
10. I avoid doing sports.	APAS	0.12	0.55	-0.35	0.26	4.06	1.18
20. I hand over strenuous activities.	APAS	0.15	0.65	0.08	-0.18	3.37	1.18
13. I laugh heartily anyway.	BES	-0.34	0.25	0.55	-0.03	3.68	1.18
16. I take it with a laugh.	BES	-0.09	0.08	0.67	0.11	2.40	1.18
17. I let my family persuade me into things, even I don't feel like it.	BES	-0.17	-0.14	0.59	-0.05	2.26	1.18
22. I distract with physical activity.	BES	0.19	-0.11	0.83	-0.02	1.75	1.18
23. I distract doing little jobs at home.	BES	0.16	0.11	0.80	0.11	2.65	1.18
4. I take care not to let myself go.	BES	-0.02	0.30	0.04	0.75	3.72	1.18
5. I try not to take any notice of it.	BES	-0.06	-0.05	0.35	0.50	3.07	1.18
6. I clench my teeth.	BES	0.08	-0.20	-0.16	0.63	3.65	1.18
11. I say "Don't make such a fuss!".	BES	-0.07	-0.07	0.19	0.62	2.80	1.18
12. I keep my appointments, even I don't feel up to it.	BES	-0.32	0.00	-0.09	0.69	4.00	1.18
15. I tell myself: "I don't have time for this right now!"	BES	0.16	-0.19	0.12	0.63	2.48	1.18
19. I carry on doing what I am doing, no matter what.	BES	-0.23	-0.19	0.04	0.49	3.46	1.18
Explained variance in percent		36.16%	11.19%	7.99%	6.63%		

Anexo 7. Cuestionario NPQ (versión española):

CUESTIONARIO DE DOLOR CERVICAL NORTHWICK PARK.

Este cuestionario va dirigido a conocer cómo puede afectar el dolor cervical a su vida diaria. Por favor, conteste cada pregunta marcando con una X, una sola alternativa.

NOMBRE _____ FECHA _____

TIEMPO DE DOLOR CERVICAL: años: _____ meses: _____ semanas: _____

1. Intensidad del dolor cervical
 - a) No tengo dolor en este momento
 - b) El dolor es leve en este momento
 - c) El dolor es moderado en este momento
 - d) El dolor es severo en este momento
 - e) El dolor es el peor imaginable en este momento
2. Dolor cervical y sueño
 - a) El dolor no me altera el sueño
 - b) El dolor ocasionalmente me altera el sueño
 - c) El dolor regularmente me altera el sueño
 - d) Duermo menos de 5 horas diarias a causa del dolor
 - e) Duermo menos de 2 horas diarias a causa del dolor
3. Pinchazos u hormigueos en los brazos por la noche
 - a) No tengo pinchazos u hormigueos por la noche
 - b) Ocasionalmente tengo pinchazos u hormigueos por la noche
 - c) Mi sueño es habitualmente alterado por pinchazos u hormigueos
 - d) A causa de los pinchazos u hormigueos duermo menos de 5 horas diarias
 - e) -A causa de los pinchazos u hormigueos duermo menos de 2 horas diarias
4. Duración de los síntomas
 - a) Mi cuello y brazos los siento normales durante todo el día
 - b) Tengo síntomas en el cuello y brazos cuando me despierto y me duran menos de 1 hora
 - c) Tengo síntomas de forma intermitente durante un tiempo al día de 1-4 horas
 - d) Tengo síntomas de forma intermitente durante un tiempo al día mayor de 4 horas
 - e) Tengo síntomas continuamente todo el día
5. Coger pesos
 - a) Puedo coger objetos pesados sin que me aumente el dolor
 - b) Puedo coger objetos pesados, pero me aumenta el dolor
 - c) El dolor me impide coger objetos pesados, pero puedo coger objetos de peso medio
 - d) Solo puedo levantar objetos de poco peso
 - e) No puedo levantar ningún peso
6. Leer y ver la T.V.
 - a. Puedo hacerlo tanto tiempo como quiero
 - b. Puedo hacerlo tanto tiempo como quiero, si estoy en una postura cómoda
 - c. Puedo hacerlo tanto tiempo como quiero, pero me produce aumento del dolor
 - d. El dolor me obliga a dejar de hacerlo m-s pronto de lo que me gustaría
 - e. El dolor me impide hacerlo
7. Trabajo
 - a. Puedo hacer mi trabajo habitual sin que aumente el dolor
 - b. Puedo hacer mi trabajo habitual, pero me aumenta el dolor
 - c. Tengo que reducir mi tiempo de trabajo habitual a la mitad por el dolor
 - d. Tengo que reducir mi tiempo de trabajo habitual a la cuarta parte por el dolor
 - e. El dolor me impide trabajar
8. Actividades sociales.
 - a. Mi vida social es normal y no me produce aumento del dolor
 - b. Mi vida social es normal, pero me aumenta el grado de dolor
 - c. El dolor ha limitado mi vida social, pero todavía soy capaz de salir de casa
 - d. El dolor ha limitado mi vida social ha permanecer en casa
 - e. No tengo vida social a causa del dolor
9. Conducir
 - a. Puedo conducir sin molestias
 - b. Puedo conducir, pero con molestias
 - c. El dolor cervical o la rigidez me limita conducir ocasionalmente
 - d. El dolor cervical o la rigidez me limita conducir frecuentemente
 - e. No puedo conducir debido a los síntomas en el cuello
10. Comparado con la ultima vez que contesto este cuestionario, su dolor de cuello esta:
 - a. Mucho mejor
 - b. Algo mejor
 - c. Igual
 - d. Algo peor
 - e. Mucho peor

Anexo 8. Cuestionario NDI (versión española):

ANEXO 1

Índice de Discapacidad Cervical

Nombre:
Fecha:
Domicilio:
Profesión:
Edad:

Por favor, lea atentamente las instrucciones:

Este cuestionario se ha diseñado para dar información a su médico sobre cómo le afecta a su vida diaria el dolor de cuello. Por favor, rellene todas las preguntas posibles y marque en cada una SÓLO LA RESPUESTA QUE MÁS SE APROXIME A SU CASO. Aunque en alguna pregunta se pueda aplicar a su caso más de una respuesta, marque sólo la que represente mejor su problema.

Pregunta I: Intensidad del dolor de cuello

- No tengo dolor en este momento
- El dolor es muy leve en este momento
- El dolor es moderado en este momento
- El dolor es fuerte en este momento
- El dolor es muy fuerte en este momento
- En este momento el dolor es el peor que uno se puede imaginar

Pregunta II: Cuidados personales (lavarse, vestirse, etc.)

- Puedo cuidarme con normalidad sin que me aumente el dolor
- Puedo cuidarme con normalidad, pero esto me aumenta el dolor
- Cuidarme me duele de forma que tengo que hacerlo despacio y con cuidado
- Aunque necesito alguna ayuda, me las arreglo para casi todos mis cuidados
- Todos los días necesito ayuda para la mayor parte de mis cuidados
- No puedo vestirme, me lavo con dificultad y me quedo en la cama

Pregunta III: Levantar pesos

- Puedo levantar objetos pesados sin aumento del dolor
- Puedo levantar objetos pesados, pero me aumenta el dolor
- El dolor me impide levantar objetos pesados del suelo, pero lo puedo hacer si están colocados en un sitio fácil como, por ejemplo, en una mesa
- El dolor me impide levantar objetos pesados del suelo, pero puedo levantar objetos medianos o ligeros si están colocados en un sitio fácil
- Sólo puedo levantar objetos muy ligeros
- No puedo levantar ni llevar ningún tipo de peso

Pregunta IV: Lectura

- Puedo leer todo lo que quiera sin que me duela el cuello
- Puedo leer todo lo que quiera con un dolor leve en el cuello
- Puedo leer todo lo que quiera con un dolor moderado en el cuello
- No puedo leer todo lo que quiero debido a un dolor moderado en el cuello
- Apenas puedo leer por el gran dolor que me produce en el cuello
- No puedo leer nada en absoluto

Pregunta V: Dolor de cabeza

- No tengo ningún dolor de cabeza
- A veces tengo un pequeño dolor de cabeza
- A veces tengo un dolor moderado de cabeza
- Con frecuencia tengo un dolor moderado de cabeza
- Con frecuencia tengo un dolor fuerte de cabeza
- Tengo dolor de cabeza casi continuo

Pregunta VI: Concentrarse en algo

- Me concentro totalmente en algo cuando quiero sin dificultad
- Me concentro totalmente en algo cuando quiero con alguna dificultad
- Tengo alguna dificultad para concentrarme cuando quiero
- Tengo bastante dificultad para concentrarme cuando quiero
- Tengo mucha dificultad para concentrarme cuando quiero
- No puedo concentrarme nunca

Pregunta VII: Trabajo y actividades habituales

Pregunta VII: Trabajo*

- Puedo trabajar todo lo que quiero
- Puedo hacer mi trabajo habitual, pero no más
- Puedo hacer casi todo mi trabajo habitual, pero no más
- No puedo hacer mi trabajo habitual
- A duras penas puedo hacer algún tipo de trabajo
- No puedo trabajar en nada

Pregunta VIII: Conducción de vehículos

- Puedo conducir sin dolor de cuello
- Puedo conducir todo lo que quiero, pero con un ligero dolor de cuello
- Puedo conducir todo lo que quiero, pero con un moderado dolor de cuello
- No puedo conducir todo lo que quiero debido al dolor de cuello
- Apenas puedo conducir debido al intenso dolor de cuello
- No puedo conducir nada por el dolor de cuello

Pregunta IX: Sueño

- No tengo ningún problema para dormir
- El dolor de cuello me hace perder menos de 1 hora de sueño cada noche
Pierdo menos de 1 hora de sueño cada noche por el dolor de cuello*
- El dolor de cuello me hace perder de 1 a 2 horas de sueño cada noche
Pierdo de 1 a 2 horas de sueño cada noche por el dolor de cuello*
- El dolor de cuello me hace perder de 2 a 3 horas de sueño cada noche
Pierdo de 2 a 3 horas de sueño cada noche por el dolor de cuello*
- El dolor de cuello me hace perder de 3 a 5 horas de sueño cada noche
Pierdo de 3 a 5 horas de sueño cada noche por el dolor de cuello*
- El dolor de cuello me hace perder de 5 a 7 horas de sueño cada noche
Pierdo de 5 a 7 horas de sueño cada noche por el dolor de cuello*

Pregunta X: Actividades de ocio

- Puedo hacer todas mis actividades de ocio sin dolor de cuello
- Puedo hacer todas mis actividades de ocio con algún dolor de cuello
- No puedo hacer algunas de mis actividades de ocio por el dolor de cuello
- Sólo puedo hacer unas pocas actividades de ocio por el dolor del cuello
- Apenas puedo hacer las cosas que me gustan debido al dolor del cuello
- No puedo realizar ninguna actividad de ocio

*Texto utilizado previamente a los cambios propuestos a raíz de los problemas de comprensión.

Anexo 9. Tabla cuestionario COM

Tabla 1
Desglose por dimensiones del CORE-OM (Evans et al. 1998)

Ítem N°	Dimensiones	Ítems*	Intensidad del ítem
4	Bienestar Subjetivo (W)	Me he sentido bien conmigo mismo**	Baja
14	Bienestar Subjetivo (W)	He tenido ganas de llorar	Alta
17	Bienestar Subjetivo (W)	Me he sentido agobiado por mis problemas	Baja
31	Bienestar Subjetivo (W)	Me he sentido optimista sobre mi futuro**	Alta
2	Problemas-Síntomas Ansiedad (P)	Me he sentido tenso, ansioso o nervioso	Baja
11	Problemas-Síntomas Ansiedad (P)	La tensión y la ansiedad me han impedido hacer cosas importantes	Alta
15	Problemas-Síntomas Ansiedad (P)	He sentido pánico o terror	Alta
20	Problemas-Síntomas Ansiedad (P)	Me ha sido imposible dejar a un lado mis problemas	Baja
5	Problemas-Síntomas Depresión (P)	Me he sentido totalmente sin energía ni entusiasmo	Alta

Ítem N°	Dimensiones	Ítems*	Intensidad del ítem
16	Riesgo Daño a sí mismo (R)	He hecho planes para acabar con mi vida	Alta
24	Riesgo Daño a sí mismo (R)	He pensado que sería mejor que estuviera muerto	Baja
34	Riesgo Daño a sí mismo (R)	Me he hecho daño físicamente o he puesto en peligro mi salud	Alta
6	Riesgo Daño a otros (R)	He sido violento físicamente hacia los demás	Alta
22	Riesgo Daño a otros (R)	He amenazado o intimidado a otra persona	Alta

Notas.
(*) Los ítems se han descrito en español de acuerdo a la traducción realizada en este estudio de adaptación;
(**) Ítems redactados en forma positiva.

Ítem N°	Dimensiones	Ítems*	Intensidad del ítem
23	Problemas-Síntomas Depresión (P)	Me he sentido desesperado o sin esperanza	Alta
27	Problemas-Síntomas Depresión (P)	Me he sentido infeliz	Baja
30	Problemas-Síntomas Depresión (P)	He pensado que yo tengo la culpa de mis problemas y dificultades	Baja
8	Problemas-Síntomas físicos (P)	He tenido molestias, dolores y otros problemas físicos	Baja
18	Problemas-Síntomas físicos (P)	He tenido dificultad para conciliar el sueño o permanecer dormido	Baja
13	Problemas-Síntomas Trauma (P)	Me han inquietado pensamientos y sentimientos no deseados	Alta
28	Problemas-Síntomas Trauma (P)	Me han angustiado imágenes o recuerdos no deseados	Alta
7	Funcionamiento General (F)	Me he sentido capaz de afrontar las cosas cuando han ido mal**	Alta
12	Funcionamiento General (F)	Me he sentido satisfecho con las cosas que he hecho**	Baja
21	Funcionamiento General (F)	He sido capaz de hacer la mayoría de las cosas que tenía que hacer**	Baja
32	Funcionamiento General (F)	He conseguido las cosas que quería**	Alta
1	Funcionamiento Relaciones Íntimas (F)	Me he sentido muy solo y aislado	Alta
3	Funcionamiento Relaciones Íntimas (F)	He sentido que tenía a alguien en quien apoyarme cuando lo he necesitado**	Baja
19	Funcionamiento Relaciones Íntimas (F)	He sentido afecto o cariño por alguien**	Baja
26	Funcionamiento Relaciones Íntimas (F)	He pensado que no tengo amigos	Alta
10	Funcionamiento Relaciones Sociales (F)	Me ha costado mucho hablar con la gente	Alta
25	Funcionamiento Relaciones Sociales (F)	Me he sentido criticado por los demás	Baja
29	Funcionamiento Relaciones Sociales (F)	Me he sentido irritable cuando estaba con otras personas	Baja
33	Funcionamiento Relaciones Sociales (F)	Me he sentido humillado o avergonzado por otras personas	Alta
9	Riesgo Daño a sí mismo (R)	He pensado en hacerme daño a mí mismo	Baja

Anexo 10. Tabla cuestionario CF-PDI

Table 3. Items of CF-PDI distribution and factor loadings according to principal component analysis with Oblimin rotation including Kaiser correction (N = 192).

	Factor 1	Factor 2
1 ¿Presenta dolor en la cara? Do you feel any pain in your face?	0.79	0.45
2 ¿Se ha visto afectada su calidad de vida por esta dolencia? Is your quality of life affected by this pain?	0.58	0.21
3 Intensidad de dolor en la cara. Pain intensity on your face.	0.68	0.23
4 Le incapacita su dolor a la hora de tener relaciones afectuosas del tipo: besos, abrazos, relaciones sexuales... Does the pain make you unable to have emotional relationships, such as: kisses, embraces, or sexual relationships?	0.69	0.06
5 ¿Tiene dolor al reír? Do you feel any pain when you laugh?	0.68	0.19
6 ¿Su dolencia hace que evite el sonreír, hablar o masticar? Does your condition make you avoiding smiling, talking or chewing?	0.44	0.13
7. ¿Tiene dolor en la mandíbula? Do you feel any pain in your jaw?	0.53	0.38
8 ¿Escucha algún ruido al mover la mandíbula? Do you hear any noise when you move your jaw?	0.40	0.23
9. ¿Nota que su mandíbula se le sale o se le traba? Do you feel your jaw getting out of place or getting stuck?	0.33	0.31
10. Intensidad de dolor al masticar Pain intensity when chewing	0.47	0.72
11. ¿Siente cansancio en la mandíbula, al hablar o al comer? Do you feel any tiredness in your jaw when you talk or eat?	0.38	0.73
12. ¿Tiene dificultad para abrir la boca? Do you have any trouble when you open your mouth?	0.23	0.73
13. Intensidad de dolor al hablar Pain intensity when talking.	0.40	0.74
14. ¿Tiene miedo de mover la mandíbula? Do you fear moving your jaw?	0.20	0.73
15. Alimentación. Nutrition	0.24	0.72
16. ¿Con qué frecuencia tiene dolor en el cuello? How often have you got any neck pain?	0.76	0.31
17. ¿Con qué frecuencia tiene dolor de cabeza? How often do you have a headache?	0.61	0.41
18. ¿Con qué frecuencia tiene dolor de oído? How often do you have an earache?	0.47	0.34
19. ¿Qué siente al tocarse la zona dolorosa? What do you feel when you touch the painful area?	0.53	0.23
20 ¿Su dolor le altera el sueño? Does the pain disrupts your sleep?	0.59	0.21
21 ¿El dolor le interfiere a la hora de desempeñar su actividad laboral? Does the pain interfere in your work?	0.38	0.36

Anexo 11. Cuestionario PCS

Patrón matricial de la Pain Catastrophizing Scale (n = 230)

Ítems	Componentes			Correlación ítem-escala
	1	2	3	
<i>Rumiación (variancia explicada: 39%)</i>				
8. Deseo desesperadamente que desaparezca el dolor	0,78	0,06	0,18	0,61
9. No puedo apartar el dolor de mi mente	0,82	0,03	0,10	0,68
10. No dejo de pensar en lo mucho que me duele	0,71	0,01	0,15	0,67
11. No dejo de pensar en lo mucho que deseo que desaparezca el dolor	0,74	0,02	0,01	0,70
<i>Magnificación (variancia explicada: 11%)</i>				
6. Temo que el dolor empeore	0,18	0,88	0,09	0,45
7. No dejo de pensar en otras situaciones en las que experimento dolor	0,04	0,72	0,03	0,31
13. Me pregunto si me puede pasar algo grave	0,11	0,69	0,10	0,38
<i>Desesperanza (variancia explicada: 10%)</i>				
1. Estoy preocupado todo el tiempo pensando en si el dolor desaparecerá	0,08	0,10	0,84	0,41
2. Siento que ya no puedo más	0,15	0,11	0,81	0,47
3. Es terrible y pienso que esto nunca va a mejorar	0,03	0,14	0,79	0,50
4. Es horrible y siento que esto es más fuerte que yo	0,28	0,02	0,81	0,57
5. Siento que no puedo soportarlo más	0,31	0,03	0,68	0,52
12. No hay nada que pueda hacer para aliviar la intensidad del dolor	0,24	0,21	0,76	0,49

Componentes: 1, rumiación; 2, magnificación; 3, desesperanza. Estadígrafos de ajuste: $\chi^2/\text{grados de libertad} = 3,0$; RMSEA (*root-mean-square error of approximation*) = 0,08; índice de ajuste comparativo = 0,91.

Anexo 12. Cuestionario SFI-SP

<u>SPINE FUNCTIONAL INDEX (SFI)</u>		FECHA: _____
NOMBRE: _____		LESIÓN: _____ <input type="checkbox"/> Cervicales <input type="checkbox"/> Dorsales <input type="checkbox"/> Lumbares
<p>POR FAVOR, COMPLETE: Su columna puede hacer que sea difícil realizar algunas cosas que normalmente hace. Esta lista contiene frases que la gente usa para describirse a sí mismos con este tipo de problemas. Piense en usted durante los últimos días. Si hay una pregunta que le describe, márkela. Sino, déjela en blanco. Si una casilla le describe sólo en parte, marque la mitad de esa casilla.</p>		
SOBRE SU COLUMNA:		
<input type="checkbox"/> 1. Me quedo en casa la mayor parte del tiempo.		
<input type="checkbox"/> 2. Cambio frecuentemente de postura para aliviar el dolor.		
<input type="checkbox"/> 3. Evito hacer trabajos pesados (Ejemplo: limpiar, levantar más de 5 kg, trabajar el jardín, etc).		
<input type="checkbox"/> 4. paro a descansar más a menudo.		
<input type="checkbox"/> 5. Pido a los demás que hagan las cosas por mi.		
<input type="checkbox"/> 6. Tengo dolor/problema casi todo el tiempo.		
<input type="checkbox"/> 7. Tengo dificultad para levantar y cargar peso (ejemplo: bolsas, compras de hasta 5 kg, etc.).		
<input type="checkbox"/> 8. Mi apetito es diferente.		
<input type="checkbox"/> 9. El caminar o el hacer mis actividades deportivas y recreativas están afectadas.		
<input type="checkbox"/> 10. Tengo dificultad con las tareas normales de la casa y la familia.		
<input type="checkbox"/> 11. Duermo peor.		
<input type="checkbox"/> 12. Necesito ayuda con mi cuidado personal (ejemplo: la ducha y la higiene)		
<input type="checkbox"/> 13. Mis actividades normales diarias (trabajo, actividades sociales) están afectadas.		
<input type="checkbox"/> 14. Estoy más irritable y/o de peor humor.		
<input type="checkbox"/> 15. Me siento débil y/o rígido.		
<input type="checkbox"/> 16. Mi independencia en el transporte está afectada (conducir, transporte público).		
<input type="checkbox"/> 17. Necesito ayuda o soy más lento/a al vestirme.		
<input type="checkbox"/> 18. Tengo dificultad para moverme en la cama.		
<input type="checkbox"/> 19. Tengo dificultad para concentrarme y/o para leer.		
<input type="checkbox"/> 20. Sentado no me encuentro cómodo.		
<input type="checkbox"/> 21. Tengo dificultad para sentarme y levantarme de la silla.		
<input type="checkbox"/> 22. Sólo me quedo de pie durante cortos periodos de tiempo.		
<input type="checkbox"/> 23. Tengo dificultad para ponerme en cuclillas o de rodillas.		
<input type="checkbox"/> 24. Tengo problemas para agacharme (ejemplo: para coger cosas, ponerme los calcetines).		
<input type="checkbox"/> 25. Uso las escaleras más lentamente o uso el pasamanos.		
SFI Puntuación: Para puntuar la parte superior, sumar las casillas marcadas.		
_____ TOTAL (SFI Puntos)	100 Escala: $100 - (\text{TOTAL} \times 4) =$ _____ %	
<p>MDC (90% CI): Cervicales = 6.9% o 1.7 SFI puntos; Dorsales y lumbares = 5.9% o 1.5 SFI puntos; Toda la columna = 6.5 % o 1.6 SFI puntos. Una puntuación menor a éstas puede ser debido a error.</p>		
Spanish translation courtesy of Dr. Cuesta-Vargas, Faculty of Health Sciences, University of Malaga, Spain		

Figure 2 Spanish Spine Functional Index.

Anexo 13. Cuestionario OREBRO

Additional file 1

ÖREBRO MUSCULOSKELETAL PAIN QUESTIONNAIRE
Spanish Version

1. Nombre:

Teléfono:

Fecha:

2. Fecha de la lesión:

Fecha de nacimiento:

3. Sexo:

4. Dónde nació:

Estas preguntas y afirmaciones son aplicables si tiene molestias o dolores como en el cuello, la espalda, el hombro.

Por favor lea y responda las preguntas cuidadosamente. No tome demasiado tiempo para responder a las preguntas, sin embargo, es importante que las conteste todas.

Siempre encontrará una respuesta para su situación particular.

5. ¿Dónde le duele? Marque la(s) casilla(s) correspondiente(s)

- Cuello Hombro Brazo Parte alta de la espalda
 Parte Baja Pierna Otro (Indicar): _____

2 x (Max 10)

6. ¿Cuántos días de trabajo ha perdido a causa del dolor durante los últimos 18 meses? Marque sólo una.

- 0 Días (1) 1 - 2 Días (2) 3 - 7 Días (3) 8 - 14 Días (4)
 15 - 30 Días (5) 1 Mes (6) 2 Mes (7)
 3 - 6 Meses (8) 6 - 12 Meses (9) Más de un año (10)

7. ¿Desde hace cuanto tiempo sufre su dolor actual? Marque sólo una.

- 0 Días (1) 1 - 2 Días (2) 3 - 7 Días (3) 8 - 14 Días (4)
 15 - 30 Días (5) 1 Mes (6) 2 Mes (7)
 3 - 6 Meses (8) 6 - 12 Meses (9) Más de un año (10)

8. ¿Es su trabajo pesado o monótono? Marque la mejor opción.

0 1 2 3 4 5 6 7 8 9 10

No en absoluto

Extremadamente

9. ¿Cómo calificaría usted el dolor que tuvo durante la semana pasada? Marque uno.

0 1 2 3 4 5 6 7 8 9 10

No dolor

Peor dolor soportable.

10. En los últimos tres meses, de media, ¿Cómo de malo fue su dolor en una escala de 0 – 10? Marque uno.

0 1 2 3 4 5 6 7 8 9 10

No dolor

Peor dolor soportable.

11. ¿Con qué frecuencia diría usted que ha experimentado episodios de dolor, de media, durante los últimos tres meses? Marque uno

0 1 2 3 4 5 6 7 8 9 10

Nunca

Siempre.

12. Sobre la base de todas las cosas para hacer frente, o lidiar con su dolor, durante un día normal ¿Cuánto es capaz de disminuirlo? Marque uno.

0 1 2 3 4 5 6 7 8 9 10

No consigo disminuirlo

Consigo eliminarlo por completo

13) ¿Cómo de tenso(a) o ansioso(a) se ha sentido durante la última semana? Marque uno.

Absolutamente calmado y relajado

Tenso y nervioso como nunca

14. ¿Cuánto se ha sentido molesto(a), por una sensación depresiva, durante la última semana? Marque uno.

0 1 2 3 4 5 6 7 8 9 10

Nunca

Extremadamente

15. En su opinión, ¿Cómo de grande es el riesgo de que su dolor actual puede llegar a ser persistente? Marque uno

16) En su opinión, ¿Cuáles son las probabilidades de que usted será capaz de trabajar en seis meses?

Ningún riesgo

Riesgo muy alto

27

0 1 2 3 4 5 6 7 8 9 10
Ninguna probabilidad Muchas posibilidades

17. Considerando sus rutinas de trabajo, gestión, salario, posibilidades de promoción y compañeros de trabajo, ¿Cual es su grado de satisfacción con su trabajo?

0 1 2 3 4 5 6 7 8 9 10
Absolutamente insatisfecho Completamente satisfecho

Éstos son algunas de las cosas que otras personas nos han dicho acerca de su dolor. Para cada afirmacion, marque un número de 0 a 10 para decir cuánto afectaría a su dolor las actividades físicas, tales como agacharse, levantar objetos, caminar o conducir.

18. La actividad física hace que mi dolor empeore.

0 1 2 3 4 5 6 7 8 9 10
Completamente en desacuerdo Completamente de acuerdo

19. Un aumento del dolor es una indicación de que debo dejar lo que estoy haciendo hasta que disminuye o desaparece.

0 1 2 3 4 5 6 7 8 9 10
Completamente en desacuerdo Completamente de acuerdo

20. No debería realizar mi trabajo habitual con mi dolor actual.

0 1 2 3 4 5 6 7 8 9 10
Completamente en desacuerdo Completamente de acuerdo

Le indicamos una lista con cinco actividades. Marque el número que mejor describe su capacidad actual de participar en cada una de estas actividades.

21. Puedo realizar una actividad ligera durante una hora.

0 1 2 3 4 5 6 7 8 9 10

22. Puedo caminar durante una hora.

0 1 2 3 4 5 6 7 8 9 10

$\frac{10-x}{}$

No puedo hacerlo debido a mi Dolor

El dolor no es un problema

23. Puedo hacer las tareas domésticas.

0 1 2 3 4 5 6 7 8 9 10

$\frac{10-x}{}$

No puedo hacerlo debido a mi Dolor

El dolor no es un problema

24. Puedo hacer la compra semanal.

0 1 2 3 4 5 6 7 8 9 10

$\frac{10-x}{}$

No puedo hacerlo debido a mi Dolor

El dolor no es un problema

25. Puedo dormir por la noche.

0 1 2 3 4 5 6 7 8 9 10

$\frac{10-x}{}$

No puedo hacerlo debido a mi Dolor

El dolor no es un problema
