



## UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MÉDICAS, ODONTOLÓGICAS Y DE LA SALUD

HOSPITAL GENERAL DE MÉXICO “DR. EDUARDO LICEAGA”

CAMPO DEL CONOCIMIENTO EN CIENCIAS MÉDICAS

### **Validación de una versión en español del instrumento *Knee injury and Osteoarthritis Outcome Score* de 12 ítems (KOOS-12) para la evaluación de pacientes mexicanos con osteoartritis primaria de la rodilla**

TESIS

QUE PARA OPTAR POR EL GRADO DE  
DOCTOR EN CIENCIAS MÉDICAS.

PRESENTA:

M en C. GABRIEL JOSE HORTA BAAS.

#### **TUTOR:**

D en C. Gloria Eugenia Queipo García.  
Hospital General de México “Dr. Eduardo Liceaga”.

D en I. María del Socorro Romero Figueroa.  
Coordinador en Investigación en Salud.  
Instituto Mexicano del Seguro Social.

#### **COMITÉ TUTORIAL**

D en C. Ingris del Pilar Peláez Ballestas  
Hospital General de México “Dr. Eduardo Liceaga”.

D en C. David Rojano Mejía.  
Jefe de la división de investigación en Salud de la UMAE Hospital de Traumatología,  
Ortopedia y Rehabilitación “Dr. Victorio de la Fuente Narvárez”.

Ciudad Universitaria, Ciudad de México. Mayo 2021.



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## Contenido

Resumen .....	4
Abstract.....	6
Marco teórico.....	8
Osteoartritis.....	8
Epidemiología .....	8
Osteoartritis de la rodilla (gonartrosis).....	11
Manifestaciones clínicas. ....	11
Clinimetría .....	12
Métodos psicométricos .....	15
Psicometría moderna.....	17
Propiedades de los instrumentos de evaluación .....	18
Teoría de respuesta al ítem.....	23
Evaluación del paciente con OA de rodilla.....	25
Justificación .....	30
Planteamiento del problema .....	32
Objetivo general .....	33
Objetivos específicos .....	33
Hipótesis .....	33
Material y Métodos.....	36
Diseño del estudio.....	36
Lugar y tiempo .....	36
Población y muestra .....	36
Sujetos .....	36
Tamaño mínimo de la muestra .....	37
Técnica muestral .....	37
Elegibilidad de los sujetos .....	37
Variables .....	38
Instrumentos.....	39
Análisis estadístico.....	43

Aspectos éticos .....	45
Resultados.....	47
Características generales de los participantes .....	47
Confiabilidad.....	48
Validez basada en la estructura interna .....	49
Modelos de la Teoría de respuesta al ítem .....	51
Validez basada en la relación con otras variables .....	58
Discusión .....	67
Conclusión .....	72
Referencias .....	73
Anexos .....	77
Anexo 1. Encuesta KOOS-12 de evaluación de la rodilla.....	78
Anexo 2. Carta de aprobación del comité local de ética .....	81

## Resumen

**Objetivo.** Evaluar la validez, confiabilidad y sensibilidad al cambio de una versión en español del instrumento *Knee injury and Osteoarthritis Outcome Score* de 12 ítems (KOOS-12).

**Métodos.** Este estudio empleó un diseño para la validación de cuestionarios. Para evaluar la validez y confiabilidad del KOOS-12 se estudiaron 199 pacientes con osteoartritis de rodilla y 10 controles sanos. La estabilidad temporal del instrumento (confiabilidad test-retest) se evaluó en 116 pacientes. La sensibilidad al cambio se evaluó en 38 pacientes que recibieron tratamiento intra-articular. La validez estructural se evaluó mediante el análisis factorial confirmatorio (AFC). Se utilizaron métodos basados en la teoría de la respuesta al ítem (TRI) para determinar el desempeño de los ítems.

**Resultados.** La consistencia interna fue adecuada para todas las escalas (alfa de Cronbach = 0.85-0.94). El coeficiente de correlación intraclase de las escalas del KOOS-12 osciló entre 0.60 y 0.71. El AFC y el Modelo de Crédito Parcial Generalizado mostraron que las escalas del KOOS-12 presentaban un buen ajuste general del modelo. No se encontró un funcionamiento diferencial de los ítems. La validez convergente se demostró por las fuertes correlaciones (rho de Spearman  $\geq 0.70$ ) con el KOOS, International Knee Documentation Committee subjective knee evaluation form (IKDC) and Knee Intermittent and Constant Osteoarthritis Pain (ICOAP). La validez de grupos conocidos mostró que el KOOS-12 discriminaba bien a los subgrupos de pacientes (daño radiográfico y estado nutricional). El valor de la respuesta media estandarizada de las escalas del KOOS-12 fueron  $\geq 0.75$ . Los

cambios en las escalas del KOOS-12 tuvieron una correlación de moderada a fuerte ( $r$  de Pearson  $\geq 0.40$ ) con los cambios en las escalas de la KOOS, ICOAP e IKDC.

**Conclusión.** La versión en español del KOOS-12 es un instrumento válido, confiable y sensible al cambio para medir la opinión sobre su rodilla y sus problemas asociados en pacientes mexicanos con osteoartritis de rodilla.

## Abstract

**Objective.** To evaluate the validity, reliability, and responsiveness to change of the 12-item Knee injury and Osteoarthritis Outcome Score (KOOS) Spanish version questionnaire.

**Methods.** This study was based on a questionnaire validation design. A cross-sectional survey of 199 patients with knee osteoarthritis (KOA) and ten healthy controls was studied to evaluate the validity and reliability of KOOS-12. One hundred and sixteen patients were assessed for test-retest reliability, and 38 patients were included for a responsiveness assessment. Structural validity was assessed by confirmatory factor analysis (CFA). Item response theory-based methods were used to determine the performance of the items.

**Results.** Internal consistency reliability was appropriate for all scales (Cronbach's  $\alpha=0.85-0.94$ ). The intra-class correlation coefficient of KOOS-12 scales ranged from 0.60 to 0.71. CFA and Generalized Partial Credit Model showed that KOOS-12 scales presented a good overall model fit. No differential item functioning was found. Convergent validity was demonstrated by strong correlations (Spearman's  $\rho \geq 0.70$ ) with KOOS, International Knee Documentation Committee subjective knee evaluation form (IKDC) and Knee Intermittent and Constant Osteoarthritis Pain (ICOAP). Known-groups validity showed that KOOS-12 well discriminated subgroups of patients (radiographic severity and nutritional status). Standardized response means for KOOS-12 scales were  $\geq 0.75$ . Changes in KOOS-12 scales had a moderate to strong correlation (Pearson's  $r \geq 0.40$ ) with the changes in the KOOS, ICOAP, and IKDC scales.

**Conclusion.** KOOS-12 Spanish version is a valid, reliable, and responsiveness to change questionnaire to measure patients' opinions about their knee and associated problems in Mexican subjects with KOA.



## **Marco teórico**

### **Osteoartritis**

El grupo de trabajo de la sociedad internacional de investigación en osteoartritis (OARSI, por sus siglas en inglés) define a la osteoartritis (OA) como una afección de las articulaciones, caracterizada por estrés celular y degradación de la matriz extracelular, iniciada por micro y macrotraumatismos que activan respuestas de reparación inadecuadas con inclusión de las vías pro-inflamatorias de la inmunidad innata. La enfermedad comienza con anomalías moleculares (metabolismo anormal de los tejidos articulares) seguidas de alteraciones anatómicas y/o fisiológicas (caracterizadas por degradación del cartílago, remodelación ósea, formación de osteofitos, inflamación articular y pérdida de la función articular normal)(1, 2).

#### *Epidemiología*

La OA es la enfermedad reumática más frecuente a nivel mundial y constituye una de las principales causas de dolor articular y discapacidad en la población adulta. Se estima que más del 80% de las personas mayores de 60 años tienen evidencia radiológica de daño articular a consecuencia de la OA(3). La OA de la rodilla y de la cadera son la causas más frecuente de discapacidad para la deambulación en las personas mayores, en todas las razas y zonas geográficas. Esta enfermedad limita la capacidad funcional(4) y disminuye de manera importante la calidad de vida de los pacientes que la sufren(5). Las deficiencias producidas por la OA (por ejemplo, el dolor y rigidez) producen discapacidad al limitar la ejecución de las actividades diarias, lo que obstaculiza los roles y la participación de los individuos en la sociedad(6). En un estudio realizado en 3 países de Latinoamérica

(Argentina, Brasil y México) del 62 al 84% de los sujetos evaluados refirieron algún grado de discapacidad, siendo los sujetos mexicanos los que reportaron la mayor prevalencia de discapacidad autopercebida(7). Otro estudio realizado en México reportó que la mayoría de los sujetos con gonartrosis presentan una calidad de vida de regular a mala con base en la escala SF-36(8). En un estudio que incluyó a 7 077 sujetos mexicanos con OA, edad media de 74 años, reportó una disminución de la capacidad funcional (HAQ-DI 0.40) y una calidad de vida medida por Euro-QoL5D de 0.69(9).

La prevalencia de la osteoartritis en México es del 10.5% (IC 95% 10.1-10.9%) con base en los resultados de un estudio en 5 regiones del país usando el modelo COPCORD (*Community Oriented Program for the Control of Rheumatic Diseases*), que incluyó a 19 213 sujetos (mujeres: 60.3%) con un edad media de 42.8 años(10). No obstante, esta prevalencia presentó variaciones importantes en las diferentes regiones del país (tabla 1). Otros estudios realizados en distintas ciudades del país informan una prevalencia de OA del 2.3-17% en la ciudad de México(11, 12), 6.6% en personas de Chihuahua(13) y del 17.4% en Yucatán(14).

<b>Tabla 1. Prevalencia estimada de osteoartritis en 5 ciudades de México</b>					
	Ciudad de México	Chihuahua	Nuevo León	Yucatán	Sinaloa
<b>Numero de sujetos</b>	4 059	1 647	1 086	3 915	4 879
<b>Osteoartritis</b>	12.8% (11.8-13.9%)	1.9% (1.3-2.7%)	16.3% (15.2-17.3%)	2.8 % (2.3-3.3%)	1.8% (1.4-2.2%)

De acuerdo a los reportes estadísticos del Instituto Mexicano de Seguro Social (IMSS) las artropatías, siendo la más frecuente la OA, constituyen uno de los diez principales motivos de consulta en el primer nivel de atención médica y en la consulta de especialidades. En el

año 2012, las artropatías representaron la 3 causa más frecuente de consulta en medicina familiar (765 929 consultas, 3.9% de las consultas totales), para el 2013 el número de consultas de 1 vez fueron de 408 987 y subsecuentes de 1 953 009; las artropatías representaron el 4 motivo de asistencia a la consulta de especialidades (238 948 consultas en el 2012 y 911 461 en el 2013)(15, 16). En los últimos 2 informes (2014-2016) el número de consultas otorgadas en el IMSS por artropatías se incrementaron a 3.3 millones de consultas, representando el 3% de todas las consultas(17, 18).

Los sitios más afectados son las rodillas, manos y cadera, que son estructuras necesarias para realizar las actividades de la vida diaria de la persona(19). En un estudio realizado en Yucatán, la rodilla fue el sitio de afectación más frecuente (72.9%), seguido de las manos (12.5%) y la cadera (2.7%); en el 11% de los casos se presentó afectación simultánea de manos y rodillas; de los sujetos con OA de rodilla el 48% presentaron discapacidad (HAQ-DI >0.25)(14). Estos resultados son similares a lo reportado en un estudio realizado en Nuevo León(6).

En México se sabe que la OA de rodilla es un serio problema de salud y que debido al aumento de personas adultas mayores y de la obesidad, se espera un incremento de su incidencia(18). La OA de rodilla sintomática en mayores de 60 años es mayor en mujeres en comparación con los hombres. Se estima que 20% de los adultos mayores con OA de rodilla sintomática presentará un grado III o IV en la siguiente década de su vida, con una prevalencia que asciende de 10% en sujetos sin obesidad a 35% con obesidad, y que entre el 10-30% de pacientes presentará dolor intenso y limitación funcional que puede condicionar discapacidad(20).

### *Osteoartritis de la rodilla (gonartrosis)*

Existe evidencia de la influencia genética en el desarrollo de la OA. Los factores genéticos explican el 60% de las artrosis de la mano y la coxartrosis, y el 40% de la gonartrosis(21, 22). Los principales factores consistentemente asociados con OA de la rodilla son la obesidad (OR 2.63, IC 95% 2.28 a 3.05), el traumatismo en la región anterior de la rodilla (OR 3.86, IC 95%: 2.61 a 5.70), la coexistencia con la OA de las manos (OR 1.49, IC 95% 1.05-2.10), el sexo femenino (OR 1.84 IC 95%: 1.32-2.55) y la edad avanzada(23).

La OA de rodilla se caracteriza por dolor articular, limitación funcional, crepitación y grados variables de inflamación. Es más frecuente en el sexo femenino, aunque después de la menopausia la relación hombre: mujer es similar(24). La frecuencia de la OA de rodilla aumenta con la edad, de manera que de acuerdo con los criterios radiográficos, el 30% de las personas entre 45 y 60 años de edad, y más del 80% de los mayores de 80 años, presentan osteoartritis en al menos una articulación. A nivel mundial es la cuarta causa de morbilidad en mujeres mayores de 60 años, y la octava en hombre(25).

### *Manifestaciones clínicas.*

Los síntomas característicos de la OA de rodilla son el dolor, la rigidez articular y la incapacidad funcional. El dolor es el síntoma principal y casi siempre sigue un patrón de características mecánicas(26). El dolor en las fases iniciales suele cursar con brotes de dolor autolimitados, generalmente empeora con el ejercicio, mejora con el reposo y los anti-inflamatorios; posteriormente, con el incremento en la pérdida del cartílago articular, el dolor

puede permanecer casi constante, incluso en el reposo y con una menor respuesta a los anti-inflamatorios(26).

Las características del dolor en los sujetos con OA de rodilla son:

- a) Comienzo insidioso, curso lento e intensidad leve o moderada.
- b) Aparece al iniciar los movimientos (marcha); después mejora y reaparece con el ejercicio intenso o prolongado.
- c) A medida que la enfermedad avanza, el dolor aparece cada vez más precozmente y con ejercicios más livianos, para acabar en un estadio en que cualquier movimiento, por pequeño que sea, lo desencadena.
- d) Cede o mejora con el reposo.
- e) Es muy raro que el dolor aparezca en la noche.

Las variables que se asocian a la presencia de dolor en la rodilla son: ser mujer, la obesidad mórbida, la afectación simultánea de artrosis de rodilla y cadera, la afectación bilateral de ambas rodillas y un grado IV de la clasificación de Kellgren-Lawrence en las rodillas(27).

La rigidez articular generalmente dura menos de 15 minutos y suele aparecer tras un periodo de inmovilidad prolongado, pero mejora rápidamente con el movimiento.

### **Clinimetria**

El término clinimetría (*clinimetrics*) se le atribuye a Alvan R. Feinstein, quién lo introdujo en la década de los ochentas para referirse a la construcción de una serie de índices, dominios o instrumentos utilizados para describir o medir síntomas, signos físicos y otros fenómenos clínicos que no se pueden medir en los estudios de laboratorio(28). Por mucho tiempo, se pensó que los síntomas y signos de las enfermedades no eran medibles o que eran subjetivos

y se confiaba en la percepción de los clínicos para estimar su presencia y magnitud. Especialmente en los síntomas subjetivos como el dolor, náusea, mareo, debilidad, etc. Las limitaciones para medirlos impedían que se manejaran como datos científicos. A partir, del desarrollo de la clinimetría se propone que es posible la medición cuantitativa de los fenómenos clínicos y así poder manejarlos como datos científicos. Medir es asignar un valor a objetos o sucesos de acuerdo a ciertas reglas(29). A partir de este paradigma, todo es medible, es decir, todo se puede expresar con números y someterse a manejo matemático o estadístico(30).

Desde entonces han surgido múltiples cuestionarios y escalas que permiten estandarizar los datos clínicos para su manejo científico y realizar comparaciones entre los pacientes, o del mismo paciente en tiempos distintos, y así juzgar objetivamente si el paciente tiene una evolución hacia la mejoría o el deterioro. En la actualidad, cada día es más habitual el empleo de escalas visuales análogas, escalas numéricas, índices de pronóstico o gravedad, escalas de capacidad funcional y criterios cuantitativos para el diagnóstico en la práctica clínica diaria(30). No obstante, previo a Feinstein, ya se conocía que para medir los fenómenos psíquicos era necesario aplicar escalas cuantitativas aun a lo subjetivo, a través de la psicometría. El desarrollo de la psicometría, tuvo lugar fuera del campo clínico, principalmente para medir fenómenos psicológicos o logros educativos en las ciencias sociales o en el área educativa, por lo cual se ha considerado que la teoría psicométrica no puede ser adaptada automáticamente a la investigación clínica(31).

En los inicios de la clinimetría, el desarrollo de los índices tenía una base empírica a partir de la observación y la entrevista de los pacientes, a diferencia de la psicometría que tenía una

base teórica(29). Wright y Feinstein(32) consideraron que existen diferencias entre la clinimetría y la psicometría, siendo su diferencia fundamental la forma de selección de los reactivos a incluirse en un índice o escala. La psicometría evalúa los constructos teóricos, a partir de una construcción teórica (o concepto) se identifican las características relacionadas y se desarrolla el índice o la escala. Por el contrario, los índices clínicos son desarrollados a partir de un proceso intuitivo durante la evaluación de los pacientes(32). Otras diferencias incluyen: los métodos que se usan para indicar la eficacia de los instrumentos de medición en psicometría (validez y confiabilidad) no evalúan la características esencial en la práctica clínica (sensibilidad); en psicometría rara vez se desarrollan instrumentos para evaluar cambios (generalmente evalúan condiciones estables) en tanto que clínicamente resulta de mayor utilidad los instrumentos sensibles al cambio; y rara vez se permite referir o se pregunta a los pacientes su estado de salud o calidad de vida, o cómo evalúan su condición actual(29).

De acuerdo acuerdo a la clinimetría, un instrumento de medición debería contar con las siguientes características(29, 33):

- 1) Debe tener un propósito y justificación. Debe definirse su función en la clínica, por ejemplo, describir el estado del paciente, indicar si ha cambiado su estado, estimar un pronóstico o ser una guía de tratamiento.
- 2) Debe tener una aplicación evidente. Debe contener variables que sean fáciles de obtener en la práctica clínica habitual, de lo contrario sólo será útil en los estudios de investigación.
- 3) El formato del índice debe comprenderse fácilmente, sus principales características incluirían la simplicidad, tener un número limitado de variables y categorías por variables, y su reproducibilidad.

- 4) La “validez de apariencia” puede determinarse únicamente por el sentido común.
- 5) La “validez de contenido” es una evaluación cualitativa de que tan apropiada es, o no, la selección o la incorporación de las variables en una escala; para que se considere válido un índice no se deben omitir las variables importantes y tampoco se deben incluir las variables que no sean pertinentes. Además, puede ser necesario ponderar las variables.

No obstante de estas diferencias, Streiner(34) considera que la clinimetría puede ser englobada dentro de la psicometría, ya que el enfoque clinimétrico no es nuevo ni único, sino más bien un subconjunto de la psicometría. Sin embargo, otros autores defienden el continuar con el uso del término de clinimetría(35).

### **Métodos psicométricos**

La medición es el proceso de asignación de números a objetos a partir de unas reglas, de modo que reflejen relaciones empíricas entre los objetos. No obstante, medir no sería únicamente cuantificar, sino la representación numérica de los hechos empíricos en sentido amplio. Tradicionalmente existen cuatro tipos de escalas según el tipo de relación representada: escala nominal, ordinal, de intervalo y de razón (36). Una ventaja de los métodos psicométricos es que proporcionan una poderosa herramienta para dotar de calidad técnica a los instrumentos(37).

Una escala es el conjunto de categorías que definen a una variable; una categoría es una de las diferentes formas de expresión que la variable puede adoptar(29). Cuando se definen las categorías con las que se identifica una variable se está desarrollando una escala o



instrumento de medición. Sin embargo, estas escalas para que sean útiles deben poseer como características esenciales: validez y confiabilidad. El propósito fundamental de una escala es lograr una aproximación válida al nivel que exhibe el examinado en el constructo de interés(37). Una escala de medida debe adecuarse a una serie de axiomas de medición (modelo) que explique la bondad de la relación entre sus variables y, de ser posible, el ajuste de los datos a las predicciones del modelo.

La métrica de un instrumento hace referencia al método utilizado para asignar valores numéricos a las respuestas dadas por el individuo y al método utilizado para la construcción de las puntuaciones. Se distinguen 3 grandes grupos de instrumentos: psicométrico, econométrico y clinimétrico. Los algoritmos de puntuación de los instrumentos psicométricos y clinimétricos se basan generalmente en la suma de las respuestas de los ítems que componen la escala (con o sin ponderación). La principal diferencia entre ellos reside en el método seguido para el desarrollo de la escala: los psicométricos se construyen sobre modelos teóricos aplicando análisis estadísticos para evaluar dicho modelo, mientras que los clinimétricos se basan casi exclusivamente en la relevancia clínica(38). Los instrumentos econométricos se basan en la teoría de la toma de decisiones con el propósito de obtener valores basados en las preferencias sobre los estados de salud de los pacientes o poblaciones; estas preferencias son conocidas como utilidades, que cuando se multiplican por los periodos de tiempo correspondientes dan lugar al cálculo de años de vida ajustados por calidad(38).

El constructo es el rango de conceptos o dominios medidos por un instrumento, es decir, su objeto de medida. El modelo de medida de un instrumento es el modelo teórico que sustenta su métrica y el nivel de agregación de su puntuación o dimensionalidad. La dimensionalidad

hace referencia al número de puntuaciones generadas, distinguiéndose índice y perfiles, aunque estas dos categorías no son exclusivas.

### *Psicometría moderna*

La psicometría, como rama de la psicología, es la disciplina encargada del desarrollo de teorías, métodos y técnicas que dan apoyo a los procesos de construcción y administración de los tests(36). Proporciona las teorías, métodos y técnicas que, en última instancia, permiten describir, clasificar, diagnosticar, explicar o predecir los fenómenos psicológicos.

La psicometría moderna se basa principalmente en dos grandes teorías: la teoría clásica de los test (TCT) y la teoría de respuesta al ítem (TRI). La TCT se basa en 3 conceptos fundamentales: la puntuación verdadera, la puntuación empírica (puntuación observada) y el error de medida. La TCT se centra en el análisis de las puntuaciones obtenidas para valorar los errores cometidos en el proceso de medida indirecta de los fenómenos. La puntuación observada respondería a la relación lineal entre la puntuación verdadera y el error de medida(36). Las principales limitaciones de la TCT están relacionadas con los instrumentos y con los sujetos que se emplean para las muestras. En el caso de los instrumentos, la TCT tiene una dependencia a las puntuaciones obtenida en el instrumento de medida; por lo cual dos instrumentos que evalúan el mismo constructo, pero emplean diferente escala de medición, sus puntuaciones no son comparables directamente. Por otro lado, la dependencia a los sujetos de medición tiene relación en que la dificultad de los ítems que conforman un instrumento depende de las características de los individuos a quienes se les administra el instrumento.

Tratando de resolver estos problemas, la teoría de respuesta al ítem (TRI) plantea la existencia de una relación entre las puntuaciones en la variable latente y la probabilidad de acertar cada ítem introduciendo diferentes funciones matemáticas para modelar adecuadamente la relación. Así, una vez seleccionada la función más adecuada se construyen las curvas características de los ítems, que representan una modelización de las respuestas teniendo en cuenta la capacidad de discriminación de los ítems, su dificultad o la probabilidad de sean acertados al azar(36).

El análisis basado en la TRI tiene el objetivo de producir puntuaciones fiables que se puedan usar para hacer inferencias válidas sobre los examinados. Los diferentes tipos de modelos de la TRI representan diferentes perspectivas teóricas. Un modelo es útil en la medida en que proporciona predicciones razonables de las probabilidades de selección de una respuesta al ítem y la ubicación de la persona en la escala, simultáneamente(39).

### *Propiedades de los instrumentos de evaluación*

#### *Confiabilidad.*

La confiabilidad evalúa la precisión con la cual un instrumento mide un constructo; es decir, proporciona mediciones libres de error(39). La confiabilidad se refiere a la precisión con la que los tests llevan a cabo la medida, es la propiedad que valora la consistencia y la precisión de la medición(36). Cualquier proceso de medición conlleva inherente un cierto error de medida, un instrumento libre de error no existe. El estudio de la confiabilidad de un instrumento de medición debe permitir conocer hasta qué punto los resultados que se obtienen a partir de su aplicación están afectados por el error de medición. Si el error es

pequeño, podemos confiar en el resultado del instrumento; si el error es grande, el uso del instrumento no tiene sentido. La confiabilidad busca estimar hasta qué punto las calificaciones obtenidas con el instrumento se deben a una varianza producida por diferencias individuales reales y no a una varianza producida por errores aleatorios. La confiabilidad de un instrumento puede ser evaluada por su consistencia interna y por su estabilidad temporal.

*Estabilidad temporal.* Un instrumento de medida fiable debe proporcionar valores estables en diferentes medidas de los mismos sujetos secuencialmente obtenidas. Si se mide un fenómeno estable del mismo sujeto en dos diferentes ocasiones con el mismo test (método test-retest), el coeficiente de correlación entre sus puntuaciones será un buen indicador de esta estabilidad del test, y por lo tanto, de su fiabilidad(36). La limitación de estos métodos, es que pueden existir factores que distorsionen las puntuaciones de los sujetos en la segunda administración. Así, si los sujetos pueden recordar el contenido del test, seguramente su rendimiento se verá mejorado respecto a la primera administración. En este sentido, un factor crucial para una correcta aplicación de este método es determinar el intervalo temporal que hay que dejar entre las dos administraciones del test.

*Consistencia interna.* La consistencia interna mide la homogeneidad de los ítems de un instrumento indicando la relación entre ellos y es el método más empleado para medir la confiabilidad de un instrumento. La consistencia interna hace referencia al grado en que cada una de las partes de las que se compone el instrumento es equivalente al resto. Se refiere al grado en que cada ítem muestra una equivalencia adecuada con el resto de los ítems, es decir, que mide con el mismo grado el constructo evaluado. El coeficiente alfa ( $\alpha$ ) de Cronbach

expresa la consistencia interna de un test a partir de la covariación entre sus ítems. Cuanto más elevada sea la proporción de la covariación entre estos ítems respecto a la varianza total del test, más elevado será el valor del coeficiente alfa ( $\alpha$ ) de Cronbach, y más elevada su fiabilidad(36). Sin embargo, una de las críticas del empleo de la teoría psicométrica en el contexto clínico es la búsqueda de una alta homogeneidad; ya que las mismas propiedades que otorgan a una escala un puntaje alto en homogeneidad pueden limitar su capacidad de detectar cambios, y así disminuir su sensibilidad(32).

En la TRI, el coeficiente de fiabilidad marginal proporciona una estimación de la precisión con que mide un instrumento. El coeficiente de fiabilidad marginal se interpreta como la proporción de varianza en la puntuación observada que se debe a la puntuación verdadera(39).

### *Validez*

La validez de los instrumentos se refiere a la confianza en que las medidas se corresponden realmente con lo que se proponen medir, es decir, el grado en que la evidencia empírica y la teoría apoyan la interpretación de las puntuaciones en relación con su uso específico(36). La validez hace referencia a la proporción de varianza real que es relevante para el atributo o dominio que se pretende medir. La validación es el proceso mediante el cual se otorga validez a un instrumento de medida. En ese sentido, la validez es entendida como: el grado en que la teoría y evidencia ayudan a interpretar las puntuaciones de un instrumento diseñado para un propósito determinado.

La forma de conceptualizar la validez ha cambiado a través del tiempo. A mediados de los años ochenta, existió un cambio relevante en la forma que se conceptualiza la validez, la validez ya no se consideró una propiedad inherente al test, sino que es más bien producto del análisis de su adecuación al propósito específico al que sirven, lo que se válida son las inferencias que se realizan a partir de este. Por ello, el responsable de asegurar la validez ya no es solo el constructor del test, sino que también lo es el usuario que emplea dicho cuestionario para una finalidad determinada. En muchas ocasiones los problemas que un cuestionario posea en lo referente a la validez se deben no al diseño del cuestionario sino a la utilización que se hace de este(36).

Validez de contenido. Hace referencia a que los ítems que componen el cuestionario sean representativos de aquello que se pretende evaluar. La evidencia de la validez de contenido hace referencia a la relación que existe entre los ítems que componen el test y lo que se pretende evaluar con él; prestando atención tanto a la relevancia (que la información que se pregunta esté directamente relacionada con lo que se pretende medir) como a la representatividad (las cuestiones que se realicen deben ser una muestra adecuada de todo lo que se pretende medir) de los ítems.

La validez de constructo se refiere a que un instrumento es válido cuando representa de manera fidedigna el constructo que pretende medir, así como las relaciones esperadas entre los diferentes constructos, con base en el sustento teórico. Existen varias formas de estimar la validez de constructo, entre ellas la diferenciación entre grupos conocidos, la concurrencia

de una prueba con otra que presenta evidencia de medir el mismo atributo y la estructura interna(40).

El estudio de la validez de constructo se considera como la principal forma de validez, este proceso puede concebirse como un caso particular de la contrastación de las teorías científicas mediante el método hipotético-deductivo. El análisis de la estructura interna evalúa si los ítems se ajustan a la dimensionalidad descrita al desarrollar el instrumento; es decir, analiza si la estructura interna del instrumento se mantiene invariante. El análisis factorial es una técnica estadística que permite evaluar la estructura interna de un instrumento, definir el número de factores y qué ítems se agrupan en cada factor. Con el análisis factorial se puede demostrar si la evidencia empírica permite aceptar el modelo basado en la teoría sobre el constructo que se evalúa o si se debe rechazar y plantear un nuevo modelo que explique el constructo medido.

Los ítems que constituyen un test pueden formar escalas unidimensionales o multidimensionales. En el caso de ser unidimensional todos los ítems, aunque de distinta manera, contribuyen a la medida de un único aspecto del constructo. Por el contrario, un test puede tener una estructura interna multidimensional, es decir, que los diferentes ítems no miden un solo aspecto o dimensión sino dos o más. La técnica estadística del análisis factorial sirve para el estudio de la contribución de los diferentes ítems a un solo factor (estructura unidimensional) o a varios factores (estructura multidimensional).

Existen 2 tipos de análisis factorial: el análisis factorial exploratorio (AFE) y el análisis factorial confirmatorio (AFC). El AFE permite explorar, basado en los datos, los posibles

constructos que explican las respuestas a los ítems de un instrumento. Una vez obtenido el número de factores, dependiendo del contenido de los ítems incluidos en cada factor, a posteriori se define el nombre del constructo obtenido. Por otro lado, el AFC precisa de una teoría que explique las respuestas a los ítems y que permita establecer las especificaciones del modelo; es decir, la construcción del modelo está basada en información a priori. El AFC comprueba si la estructura factorial es consistente (presenta un buen ajuste a los datos) o no con la estructura teórica y se emplea para demostrar la validez de constructo de los modelos obtenidos en el AFE(39).

La relación con otras variables que miden el mismo constructo (validez convergente) o diferentes constructos (validez divergente) son relevantes cuando las puntuaciones del instrumento se usan para estimar el nivel de los pacientes en un constructo, dado que proporcionan información de la magnitud en que estas relaciones son acordes con el constructo en el que se basa la interpretación de los resultados. Se espera que la correlación entre los instrumentos que miden el mismo constructo sea mayor que entre los instrumentos que miden constructos diferentes. La validez de grupos conocidos se demuestra cuando las puntuaciones del instrumento discriminan entre los grupos de pacientes que, teóricamente, se espera que sean diferentes en el constructo medido(39).

#### *Teoría de respuesta al ítem*

La teoría de respuesta al ítem (TRI) se basa en el cálculo, para cada ítem, de una serie de parámetros, que asumen un modelo matemático muy concreto. El objetivo fundamental es la medición del rasgo latente, a partir de tres parámetros: discriminación del ítem, dificultad del ítem y el acierto al azar(36). Estos parámetros son representados por las curvas características



del ítem (CCI). La CCI muestra, en el eje de las ordenadas (el eje  $Y$ ), la probabilidad de acertar el ítem a partir de la magnitud del rasgo latente (eje de las abscisas o eje  $X$ ). Esta probabilidad sigue una función sigmoïdal (en forma de S, también llamada logística) y el rasgo latente se indica como  $\theta$  (*theta*). La función logística tiene como propiedad que no puede ser menor de 0 ni mayor de 1, y el valor de  $\theta$  suele estar comprendido entre -3 y +3(36).

La TRI agrupa una serie de modelos que, por medio de funciones matemáticas, describen la probabilidad que tiene un paciente de seleccionar una determinada respuesta a un ítem de acuerdo con su nivel del rasgo latente (Tabla 2). La TRI permite mediciones invariantes más allá de los ítems que componen el instrumento y provee información de la exactitud con que se mide el constructo en función del nivel del rasgo latente ( $\theta$ ). En la TRI el ítem se considera la unidad de análisis y el nivel del rasgo latente que presenta el paciente se estima a partir del patrón de respuesta obtenido del conjunto de ítems. El número de rasgos latentes que intervienen al contestar un ítem establece si el modelo es unidimensional (un rasgo latente) o multidimensional (más de un rasgo latente). Los modelos de la TRI vinculan el nivel  $\theta$  del paciente con los parámetros del ítem (discriminación y dificultad) y la probabilidad de seleccionar una respuesta. El parámetro de discriminación ( $\alpha$ ) mide la fuerza de la relación entre el ítem y el rasgo latente que se mide. Los parámetros de dificultad (parámetros  $\beta$ ) se interpretan como desviaciones estándar que muestran el rango del rasgo latente cubierto por el ítem. Cuanto mayor sea el parámetro  $\beta$  mayor será el nivel de limitación funcional que una persona debe tener para seleccionar esa opción de respuesta. Las pruebas de bondad de ajuste permiten evaluar hasta qué punto el modelo representa los datos observados. Si el modelo se ajusta a los datos empíricos, se puede suponer que el modelo representa de forma apropiada

la relación entre el rasgo latente y la probabilidad de que el paciente seleccione determinada respuesta al ítem(39).

### *Evaluación del paciente con OA de rodilla*

Las medidas de desenlace desde la perspectiva del paciente (MDPP) son instrumentos que cubren un espectro que va desde la evaluación de los síntomas, el estado funcional, el estado de salud, la calidad de vida relacionada a la salud, el bienestar o la satisfacción respecto al cuidado o tratamiento recibido, desde la perspectiva del paciente sin la interpretación de sus respuestas por parte del médico o de otros profesionales de la salud(41, 42). El objetivo de emplear estas medidas es que el propio paciente sea el evaluador de su estado de salud(41).

Las MDPP tienen interés tanto en la práctica clínica diaria como en la investigación. En el contexto de la práctica clínica diaria, el uso de las MDPP permite realizar una medicina más participativa al involucrar la perspectiva del paciente y permite tener una estimación estandarizada de la información verbalizada por la persona. A nivel de investigación clínica, disponer de MDPP mejorará la medición del impacto de la enfermedad desde la perspectiva de quien la padece. De esta manera permitirá evaluar si las nuevas intervenciones en la salud son eficaces y efectivas para el eje central de la atención clínica, el paciente. Aunado a lo anterior, se ha descrito que la evaluación de las MDPP puede estar influida por el contexto sociocultural del paciente, por la forma y el lugar de la aplicación de los cuestionarios y por la familiaridad que tengan los pacientes para contestar estos cuestionarios(42). Desde una perspectiva ética, las MDPP ofrecen un elemento para reforzar la idea de autonomía del

paciente y de la beneficencia por parte del médico a través de mejorar la comunicación y la toma de decisiones(42).

Desde el punto de vista de la teoría de la medición los principales síntomas de las personas con OA están formados por constructos (por ejemplo, el dolor, la rigidez y calidad de vida), es decir, son un concepto hipotético cuya existencia nunca podrá ser totalmente confirmada. El valor de un instrumento de medición se encuentra en su capacidad para medir el fenómeno de interés de forma precisa. Debido a que la validez no es una característica intrínseca del instrumento, el proceso de validación y estandarización requiere de un proceso iterativo. La adaptación transcultural, y no solo la traducción al idioma deseado, al medio donde se pretende usar es un paso obligado. En las diferentes naciones o poblaciones hispanohablantes pueden existir palabras de uso exclusivo o con significado diferente a los países o grupos culturales. No obstante, en general, puede no existir diferencias culturales tan marcadas, que impidan el uso de estos cuestionarios en los diferentes países que hablan español(43).

#### *Knee injury and Osteoarthritis Outcome Score*

El instrumento *Knee injury and Osteoarthritis Outcome Score* (KOOS) fue desarrollado por Roos *et al.*(44) en 1998, para evaluar pacientes jóvenes con lesiones deportivas como lesiones del ligamento cruzado anterior, lesiones meniscales y artrosis postraumática de rodilla. Se desarrolló como una extensión del índice de osteoartritis WOMAC con el propósito de evaluar los síntomas y la función a corto y largo plazo. El instrumento está formado por cinco subescalas: dolor (9 ítems); síntomas (7 ítems); actividades de la vida cotidiana (17 ítems); actividades recreativas y deportivas (5 ítems) y calidad de vida (4 ítems). Cada uno de los

ítems tiene como opción de respuesta una escala Likert donde el 0 implica sin problemas y 4 problemas graves. El tiempo de cumplimentación promedio de los 42 ítems es de aproximadamente 10 minutos. Su uso incluye a pacientes con artrosis y traumatismos de la rodilla entre los que se incluyen reparación de ligamentos cruzados y lesiones de meniscos. El KOOS es validado para evaluar intervenciones ortopédicas, como la reconstrucción del ligamento cruzado anterior, la meniscectomía y el reemplazo total de rodilla. Además, el instrumento se ha utilizado para evaluar la fisioterapia y los suplementos nutricionales(45). La amplia aceptación de este instrumento, se ve reflejada por la validación del instrumento en 35 poblaciones(46).

Confiabilidad. La estabilidad temporal con el método test-retest se demuestra con el coeficiente de correlación intraclass (ICC) de 0.78 a 0.97(47). El ICC descrito para las subescalas del KOOS son: 0.85-0.86 para el dolor, 0.84-0.93 para los síntomas, 0.75-0.91 para las actividades de la vida diaria, 0.78-0.82 para las actividades deportivas y recreativas, y 0.83-0.86 para la calidad de vida(44, 48). La consistencia interna medida por el coeficiente alfa de Cronbach para cada una las subescala es: dolor 0.84, síntomas 0.74, actividades de la vida diaria 0.95, actividades deportivas y recreativas 0.85 y calidad de vida 0.71(48).

Validez. La validez de contenido del instrumento se basó en la revisión de la literatura, entrevistas con pacientes con lesiones del menisco y del ligamento cruzado anterior y en sujetos con OA temprana, y en la evaluación por un grupo de expertos (médicos especialista en cirugía ortopédica y especialistas en terapia física)(44). La validez de constructo convergente y divergente se determinó en comparación con el SF-36 y con la escala de puntaje de rodilla de Lysholm(47, 48). El coeficiente de correlación de Spearman estimado

entre la subescala de dolor corporal del instrumento SF-36 y la subescala del dolor KOOS es del 0.46-0.62, entre las subescalas de función física y actividades de la vida diaria del 0.47-0.48 y entre la función física y actividades deportivas y recreativas del 0.47(47, 49). La evaluación de la dimensionalidad de las subescalas de la versión sueca del cuestionario KOOS, mediante el análisis de componentes principales, demostró la unidimensionalidad de las subescalas actividades de la vida diaria, actividades deportivas y recreativas y calidad de vida, pero no de las subescalas de dolor y de los síntomas (al extraerse 2 factores en cada una estas subescalas)(48).

Sensibilidad al cambio. El tamaño del efecto estimado posterior a los procedimientos quirúrgicos varió dependiendo del procedimiento y de la subescala específica del KOOS. El mayor tamaño del efecto se presentó posterior al tratamiento con el reemplazo total de rodilla. La mejoría mínima clínicamente significativa para el KOOS la representa un cambio de 8-10 puntos(45). El KOOS ha demostrado ser más sensible que el WOMAC en sujetos con OA de rodilla tratados con reemplazo articular, la subescala sensible al cambio fue la calidad de vida relacionada con un tamaño del efecto de 2.86 a los 6 meses y de 3.54 a los 12 meses. La segunda subescala más sensible fue la del dolor, con un tamaño del efecto de 2.28 y 2.55 a los 6 y 12 meses, respectivamente. La subescala de actividades deportivas y recreativas fue la menos sensible con un tamaño del efecto de 1.18 y 1.08 a los 6 y 12 meses, respectivamente(47). En otros estudios el tamaño del efecto reportados en cada subescala son: dolor 0.84, síntomas 0.87, actividades de la vida diaria 0.94, actividades deportivas y recreativas 1.16 y calidad de vida 1.65(44).

Permiso para su uso. No se requiere permiso ni pago de derechos para usar el KOOS.

### *12-item Knee injury and Osteoarthritis Outcome Score*

A pesar de las buenas propiedades psicométricas del KOOS, el tener que completar 42 ítems supone una carga importante para el paciente. Además, por su extensión se considera que requiere demasiado tiempo para su uso clínico habitual (50). Por lo cual se consideró necesario desarrollar una versión corta del instrumento pero que conservara sus propiedades psicométricas.

La versión de 12 ítems del KOOS (KOOS-12) se desarrolló empleando métodos basados en teoría de respuesta al ítem (TRI). Los ítems de la KOOS-12 se seleccionaron en función de su contenido y de la información proporcionada por los pacientes, los médicos y los investigadores sobre la importancia de los ítems y su capacidad de traducción(51). Para la selección de los ítems más apropiados se emplearon los modelos de la TRI y simulaciones de evaluaciones adaptativas computarizadas (CAT)(51). La evaluación de las propiedades psicométricas de la versión en inglés demostró que es un instrumento válido y confiable para su uso en pacientes con OA de rodilla. El KOOS-12 mostró una buena validez convergente, no presentó un efecto techo y es sensible al cambio en los pacientes sometidos a cirugía(50).

En el ámbito clínico, un instrumento corto que sea válido, confiable, sensible al cambio y que proporcione medidas específicas de dominio junto con una puntuación global del impacto en la OA rodilla, sería óptimo(50). El KOOS-12 proporciona un instrumento con estas características. Por lo cual, el KOOS-12 constituye un instrumento ideal para su uso clínico.

## **Justificación**

La medición es indispensable en medicina, ya que su resultado puede sustentar las decisiones clínicas(52). No obstante, el proceso de validación de un instrumento de medición es particularmente difícil y es a menudo una práctica esporádica en lugar de habitual; e incluso los instrumentos de medición pueden aceptarse y usarse en la práctica clínica antes de que se hayan realizado los procesos formales de validación(52).

En la medicina clínica lo subjetivo resulta particularmente valioso, pues a partir de lo que el enfermo siente se inicia la relación médico-paciente(30). La subjetividad del paciente se expresa en síntomas y en la capacidad de percibir los efectos de su enfermedad(30). En la investigación clínica, el empleo de instrumentos de medición válidos permite tener resultados objetivos cuantificables de los fenómenos clínicos subjetivos. El empleo de instrumentos válidos permite desarrollar investigaciones para describir atributos poblacionales, compararlos con otras poblaciones o con una misma población mediada por un intervalo de tiempo, e incluso valorar la efectividad de una determinada intervención.

En el caso de los pacientes con osteoartritis, el cómo se siente y funciona puede tener más relevancia que las mediciones radiográficas o los estudios de laboratorio. El dolor y la discapacidad funcional asociada a la OA son las principales razones por las que los pacientes buscan tratamiento(44). Por lo tanto, estos desenlaces deben considerarse primordiales para evaluar la respuesta a la cirugía, fisioterapia o cualquier tratamiento. Debido a que existen varios instrumentos para la evaluación de los pacientes con OA de rodilla que se utilizan en la investigación, es fundamental que las decisiones sobre cuál utilizar se basen en atributos

específicos(53). La evaluación de los instrumentos por medio del uso de métodos psicométricos permite establecer evidencias de su grado de validez y confiabilidad, para garantizar su adaptación transcultural y determinar su aplicabilidad en la práctica clínica.

Al igual que con cualquier instrumento nuevo, el KOOS-12 debe estudiarse en poblaciones distintas de los pacientes con artroplastía total de rodilla, como los pacientes con OA de leve a moderada y en países fuera de los Estados Unidos(54). Los médicos que trabajan con pacientes cuya lengua materna no es el inglés deben tener en cuenta que el contenido y las propiedades psicométricas pueden diferir de las del instrumento original. Y es indispensable comprobar si la versión en español del instrumento mantiene la validez y confiabilidad que demostró en su versión original.



## **Planteamiento del problema**

En los últimos años, la OA de rodilla adquiere mayor importancia por el incremento de su prevalencia. Dado que la intensidad del dolor, la evaluación global y la función física no pueden ser evaluadas directamente por el médico, es necesario contar con instrumentos de medición estandarizados para su evaluación. Se cuenta con instrumentos válidos y sensibles para detectar cambios desarrollados en el idioma inglés y que no cuentan con evidencia de su validez en población mexicana. Basándose en la falta de información de las propiedades psicométricas del KOOS-12 en pacientes mexicanos con OA de rodilla se plantea la siguiente pregunta de investigación:

¿La versión en español del KOOS-12 mantiene el mismo grado de validez, confiabilidad y sensibilidad al cambio que se demostró en su versión original en inglés, para la evaluación de los sujetos con osteoartritis primaria de rodilla?

## **Objetivo general**

Evaluar la validez, confiabilidad y sensibilidad al cambio de una versión en español del instrumento *Knee injury and Osteoarthritis Outcome Score* de 12 ítems (KOOS-12).

## **Objetivos específicos**

1. Describir el grado de dolor, funcional y nivel de calidad de vida de los pacientes con artrosis de rodilla.
2. Evaluar si alguno de los ítems presenta un diagnóstico diferencial del ítem.
3. Evaluar la sensibilidad al cambio posterior al ajuste de tratamiento.

## **Hipótesis**

Si la versión en español del KOOS-12 mantiene un grado de validez y confiabilidad similar al de su versión original en inglés, entonces se demostrará una adecuada confiabilidad, validez y sensibilidad del cambio.

Para la evaluación de la validez convergente se contrastaron las siguientes hipótesis:

- 1) La correlación entre la escala de dolor del KOOS-12 y la escala del dolor, actividades de la vida diaria y calidad de vida del instrumento KOOS debe ser fuerte y positiva ( $r > 0.7$ )(54).

- 2) La correlación entre la escala del dolor del KOOS-12 y el instrumento ICOAP (*Knee Intermittent and Constant Osteoarthritis Pain*) debe ser fuerte y negativa ( $r > -0.6$ )(55-57).
- 3) La escala de función del KOOS-12 debe presentar una correlación positiva y fuerte ( $r > 0.7$ ) con la escala de actividades de la vida diaria del KOOS(54).
- 4) La escala de funcionamiento del KOOS-12 presenta una correlación positiva y moderada ( $0.4 < r < 0.7$ ) con la escala de actividades deportivas del KOOS y con las pruebas de rendimiento físico (*Test up and Go* y *30-s chair stand test*).
- 5) La puntuación de las escalas de dolor y calidad de vida del KOOS-12 presentan una correlación positiva y moderada ( $0.4 < r < 0.7$ ) con la puntuación del instrumento IKDC (*International knee documentation committee subjective knee evaluation form*).
- 6) La puntuación de la escala funcionamiento del KOOS-12 presenta una correlación positiva y fuerte ( $r > 0.7$ ) con la puntuación del instrumento IKDC (*International knee documentation committee subjective knee evaluation form*).
- 7) La correlación entre la escala de funcionamiento del KOOS-12 y las escalas de movilidad y actividades de la vida diaria del WHODAS-2.0 debe ser de moderada a fuerte ( $0.4 < r < 0.8$ ).

Para evaluar la validez divergente se contrastaron las siguientes hipótesis:

- 1) La escala de dolor del KOOS-12 presenta una correlación positiva y moderada ( $0.4 < r < 0.69$ ) con las escalas de síntomas y actividades deportivas del KOOS(54).
- 2) La escala de calidad de vida del KOOS-12 presenta una correlación positiva y moderada con las escalas de dolor, síntomas y actividades deportivas del KOOS.

- 3) La correlación entre la escala de funcionamiento del KOOS-12 y la escala de discapacidad cognitiva del WHODAS 2.0 debe ser baja ( $r < 0.4$ ).

Para evaluar la validez de grupos conocidos se contrastaron las siguientes hipótesis:

- 1) La puntuación de las escalas y la puntuación del resumen del KOOS-12 en el grupo de personas sanas debe ser significativamente mayor que la puntuación de los pacientes con OA de rodilla.
- 2) La puntuación de las escalas y del resumen del KOOS-12 es significativamente mayor en los pacientes con OA de rodillas grado 1 comparado a los pacientes con OA grado 3 y 4.
- 3) La puntuación de las escalas del KOOS-12 es significativamente menor en los pacientes con OA de rodilla obesos en comparación con los pacientes no obesos.

## **Material y Métodos**

### **Diseño del estudio**

Para resolver la pregunta de investigación, se diseñó un estudio de las siguientes características:

Por la maniobra del investigador: Observacional

Por el número de mediciones: Longitudinal

Por la recolección de los datos: Prolectivo

Por la dirección: Prospectivo.

Por el diseño: Cohorte.

### **Lugar y tiempo**

El estudio se realizó en dos hospitales del Instituto Mexicano del Seguro Social (IMSS). Hospital General Regional número 220 “Gral. Vicente Villada”, Toluca, Estado de México; durante el periodo del 1 de agosto de 2017 al 15 de junio de 2018. Hospital General Regional 1 “Lic. Ignacio García Téllez”, Mérida, Yucatán; durante el periodo del 1 de julio de 2018 al 31 de diciembre de 2019.

### **Población y muestra**

#### *Sujetos*

La población está conformada por pacientes con diagnóstico de osteoartritis rodilla referidos de la consulta externa de Reumatología y Traumatología y Ortopedia.

### *Tamaño mínimo de la muestra*

El cálculo del tamaño de la muestra se basó en las recomendaciones de expertos en el tema. Para el análisis factorial, el tamaño de la muestra debe ser al menos siete veces el número de ítems (es decir, 70 pacientes) con un mínimo de 100(58). Con respecto al tamaño de la muestra para el análisis de la TRI, una muestra de  $\geq 200$  pacientes se considera el tamaño mínimo para los modelos logísticos de dos parámetros(59).

### *Técnica muestral*

Se realizó un muestreo de casos consecutivos de los pacientes que aceptaron participar en el estudio.

## **Elegibilidad de los sujetos**

### *Criterios de inclusión:*

- ⇒ Pacientes mayores de 45 años con diagnóstico clínico-radiográfico de osteoartritis de rodilla primaria.
- ⇒ Pacientes que cumplan con los criterios de clasificación del Colegio Americano de Reumatología(60) para osteoartritis primaria de rodilla.
- ⇒ Para investigar la validez de los grupos conocidos, se invitó a participar a personas sanas (familiares de los pacientes sin OA de rodilla) mayores de 18 años.

### *Criterios de exclusión:*

- ⇒ Pacientes con artritis aguda de la rodilla.
- ⇒ Pacientes con osteoartritis secundaria: Acromegalia, Hemocromatosis y Hemofilia.

- ⇒ Pacientes con coexistencia de Artritis Reumatoide, Espondiloartritis axial o periférica, Lupus Eritematoso Sistémica, Fibromialgia o alguna otra enfermedad reumática que pueda ocasionar una osteoartritis secundaria.
- ⇒ Presencia de deterioro cognitivo.
- ⇒ Pacientes con prótesis articular de rodilla.
- ⇒ Pacientes con cirugía de rodilla en los 6 meses previos a la inclusión al estudio.
- ⇒ Pacientes con diagnóstico de neuropatía diabética.

Los *criterios de eliminación* serán:

- ⇒ Pacientes que contesten menos del 80% de los cuestionarios.

## **Variables**

*Instrumentos de medición.*

Definición conceptual: Es un procedimiento de medida objetiva y estandarizada de una muestra de comportamientos.

Definición operacional: Resultado de la calificación del cuestionario KOOS-12 y de cada una de sus escalas.

Escala de medición: Cuantitativa continúa.

Indicadores: Puntuación de 0 a 100.

*Obesidad.*

Definición conceptual: Es una enfermedad sistémica, crónica, progresiva y multifactorial que se define como una acumulación anormal o excesiva de grasa.

Definición operaciones: Se definirá la presencia de obesidad de acuerdo al resultado del índice de masa corporal (IMC). De acuerdo con la clasificación de la Organización Mundial de Salud, un sujeto con  $IMC \geq 30$  se considera obeso.

Escala de Medición: Nominal.

Indicadores: Presente o Ausente

Intensidad del dolor.

Definición conceptual. Experiencia sensorial y emocional desagradable, asociada con un daño tisular real o potencial, o descrita en términos de dicho daño.

Definición operacional: Puntuación obtenida al calificar la escala de intensidad ICOAP.

Escala de Medición: Cuantitativa discreta.

Indicadores: 0 a 100.

## **Instrumentos**

Se decidió incluir instrumentos de auto-reporte, ya que pueden facilitar la recolección de datos y la libre expresión de la sintomatología de los sujetos y no está de por medio el sesgo del observador. El empleo del KOOS, KOOS-12, ICOAP, IKDC y del WHODAS 2.0 permitirá la comparación de nuestros resultados con los reportados en otras poblaciones.

**KOOS.** Consta de 42 ítems en 5 escalas: 1) dolor, mide la frecuencia y gravedad del dolor durante las actividades de la vida diaria; 2) síntomas; 3) funcionamiento en las actividades de la vida diaria (ADL), mide la dificultad experimentada durante las actividades de la vida diaria; 4) actividades deportivas y recreativas, mide la dificultad experimentada durante las



actividades deportivas y recreativas; 5) calidad de vida. Los pacientes responden a cada ítem basándose en el estado de su rodilla en la semana previa a la evaluación, empleando una escala de respuesta de 5 categorías. Estas subescalas se puntúan por separado, no se recomienda una puntuación total. Las puntuaciones se transforman en una escala de 0 a 100; las puntuaciones más altas representan mejores resultados(61).

**KOOS-12.** Está compuesto de 3 escalas específicas que miden el dolor, la función y la calidad de vida específica de la rodilla. Para calcular la puntuación de la escala se debe responder al menos a la mitad de los ítems, y se calcula una estimación específica de la persona para cualquier ítem faltante. Las puntuaciones se transforman en una puntuación de 0 a 100, siendo 0 y 100 la peor y la mejor puntuación posible, respectivamente. La puntuación de resumen del impacto de la rodilla de la KOOS-12 se calculó como la media de las puntuaciones de las tres escalas. No se calcula una puntuación de resumen si falta alguna de las tres escalas(54).

**Knee Intermittent and Constant Osteoarthritis Pain (ICOAP).** Consta de 11 ítems. Cinco ítems evalúan el dolor constante y seis ítems evalúan el dolor intermitente. Los pacientes responden a cada uno de ellos en función del estado de su rodilla durante la semana previa a la evaluación en una escala de respuesta de 5 categorías. La puntuación total se obtiene sumando las puntuaciones de los ítems y normalizando de 0 (sin dolor) a 100 (dolor extremo); las puntuaciones más altas representan un mayor grado de dolor(62).

**International knee documentation committee subjective knee evaluation form (IKDC).** Comprende 18 ítems, en los dominios de síntomas, funcionamiento durante las actividades

de la vida diaria y durante los deportes, evalúa la función actual de la rodilla y la participación en el trabajo y en los deportes. La puntuación total se calcula mediante la siguiente fórmula:  $(\text{suma de ítems})/(\text{puntuación máxima posible}) \times 100$ . La puntuación posible va de 0 a 100, donde 100 significa que no hay limitación en las actividades diarias o deportivas y la ausencia de síntomas(63).

**World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0).** Mide las limitaciones en la actividad y la restricción de las personas según los constructos incluidos en la Clasificación Internacional del Funcionamiento, la Discapacidad y la Salud (CIF). Las subescalas de cognición, movilidad y actividades de la vida diaria del WHODAS 2.0 son cuestionarios de auto-reporte de seis, cinco y ocho ítems, respectivamente. Las puntuaciones de cada una de las subescalas se crean sumando las puntuaciones de los ítems y normalizándolas de 0 (sin discapacidad) a 100 (discapacidad total)(64).

#### *Pruebas de rendimiento físico.*

**30 s-chair stand test (la prueba de levantarse de una silla en 30 segundos).** Esta prueba evalúa la fuerza de las extremidades inferiores como un indicador del estado funcional y se ha evaluado su fiabilidad y validez en sujetos con OA de rodilla(65). El test comienza con el paciente sentado en el medio de la silla, con la espalda recta y con toda la planta de los pies apoyada sobre el suelo, con los brazos cruzados por las muñecas y las muñecas sobre el pecho. A la señal de “YA” el participante se pone completamente de pie y posteriormente vuelve a colocarse en la posición de sentado. Se anima al participante a que se ponga de pie tantas veces como sea posible durante 30 segundos. Antes de la realización del test se

realizará una demostración por el propio evaluador. Por seguridad la silla debe ser colocada con el respaldo contra la pared, o de alguna manera que la estabilice, para prevenir que se mueva durante la realización del test; igualmente se suspenderá la prueba si el paciente siente que se puede caer. La puntuación es el número de veces que el sujeto es capaz de levantarse durante 30 segundos. Si el sujeto realiza más de media subida al final de los 30 segundos se considerará como una subida completa. La prueba se realizó por duplicado, anotando el número de veces en cada una de ellas para calcular la media de los 2 intentos.

**Timed Up and Go test (TUG).** La transferencia desde la posición sedente a bípedo es una tarea motora que requiere una adecuada capacidad motora y que determina el nivel de funcionalidad de una persona. La prueba evalúa la movilidad y consiste en pasar de estar sentado a estar de pie, permanecer en posición ortostática, caminar, girar y pasar de estar de pie a estar sentado. A cada persona se le pidió que tomara asiento en forma cómoda en una silla estándar sin apoyabrazos de 42 cm de altura, frente a una indicación verbal por parte del evaluador se le solicitó al sujeto que se levantara desde la silla y caminar en línea recta hasta una distancia de 3 metros marcados en el suelo, una vez que el sujeto gira y regresa a sentarse en la silla, todo realizado de la manera más rápida posible, midiendo el tiempo que se tarda en realizar estas tareas. La prueba se realizó por duplicado con un intervalo de un minuto, el puntaje consignado fue el promedio del tiempo de las 2 repeticiones(65).

#### *Evaluación del daño radiográfico.*

Para la evaluación de la rodilla se empleó la proyección antero-posterior (AP) con carga y en extensión completa(5). El daño radiográfico se evaluó mediante la escala de Kellgren y

Lawrence(49). Esta escala permite clasificar la gravedad de la osteoartritis de rodilla en una escala ordinal:

1. Grado 0 (Ausente): Corresponde a una radiografía normal.
2. Grado I (Dudosa): Disminución dudosa del espacio articular medialmente y osteofitos posibles en los bordes.
3. Grado II (Mínima): Osteofitos definitivos y disminución posible del espacio articular.
4. Grado III (Moderada): Osteofitos múltiples moderados, disminución definitiva del espacio articular y cierta esclerosis y deformidad ósea.
5. Grado IV (Grave): Grandes osteofitos, disminución acentuada del espacio articular, esclerosis subcondral intensa, y deformación ósea terminal.

### **Análisis estadístico**

El análisis descriptivo de las variables categóricas (características de los pacientes y de los ítems) y el efecto suelo-techo (pacientes que obtenían la puntuación mínima y máxima posible, respectivamente) se presentan como el número de casos (n) y su porcentaje (%). En el caso de las variables continuas se presentan como media y desviación estándar (media  $\pm$  DE) o mediana (rango intercuartil), según corresponda. Para la comparación entre las puntuaciones del factor 1 y del factor 2 se empleó la prueba T para muestras relacionadas. La concordancia entre el HAQ-DI y el HAQ-II se evaluó con el coeficiente de correlación y concordancia de Lin (CCC) y por el método de Bland-Altman(66).

*Análisis factorial.* El AFC se realizó utilizando correlaciones policóricas y el procedimiento de estimación de los parámetros empleado fue el de *mínimos cuadrados*

*ponderados diagonalmente (DWLS)(67)*. Los criterios que se emplearon para determinar un buen ajuste del modelo son: una razón de ji-cuadrado ( $\chi^2$ ) sobre los grados de libertad (CMIN/DF) < 3, el error cuadrático medio de aproximación (RMSEA) < 0.06, el *residuo cuadrático medio estandarizado* (SRMR)  $\leq$  0.08, el índice de ajuste comparativo (CFI)  $\geq$  0.95 y el índice de Tucker Lewis (TLI)  $\geq$  0.95. Un valor RMSEA  $\leq$  0.08 se considera un ajuste aceptable y un valor > 0.10 se considera un mal ajuste(68). En el AFE, el número de factores apropiado a extraer se determinó con el análisis paralelo, los parámetros se estimaron con el método de residuales mínimos (minres) y se empleó una rotación oblimin.

*Validez basada en la relación con otras variables*. Se calculó el coeficiente de correlación de Spearman ( $\rho$ ) entre las puntuaciones del HAQ-II y las puntuaciones de los otros instrumentos. Se empleó la prueba de Kruskal-Wallis y la prueba de Dunn para saber entre qué grupos las puntuaciones del HAQ-II diferían significativamente acorde al grado de actividad medido por el CDAI(69).

*Fiabilidad*. La consistencia interna se estimó con los coeficientes alfa de Cronbach y omega de McDonald. Estos coeficientes tienen la misma interpretación, un valor entre 0.70 y 0.95 se considera apropiado(70). La estabilidad temporal de las escalas del instrumento KOOS-12 se evaluaron mediante el coeficiente de correlación intraclase (CCI<sub>3,1</sub>, modelo ANOVA de dos factores, efectos mixtos, con interacción para la concordancia absoluta entre las puntuaciones individuales). Un CCI  $\geq$  0.70 se consideró adecuado(70). Para estimar el error de medición se determinó el error estándar de medida (EEM), el cambio mínimo detectable (CMD) y los límites de acuerdo de Bland-Altman(70).

*Análisis basado en la TRI.* Basándose en los estudios previos en el idioma inglés, se analizó el desempeño de los ítems del instrumento KOOS-12 con el modelo de crédito parcial generalizado (MCPG). El ajuste del modelo a los datos a nivel del ítem se evaluó con el estadístico S-X2, un valor de p ajustado mediante el método de Benjamini-Hochberg  $< 0.05$  indica un mal ajuste a nivel del ítem(68). Se consideró un buen ajuste global del modelo si el valor de p del estadístico de información limitada (M2\*) era  $> 0.05$ , RMSEA  $< 0.089$  y SRMR  $< 0.05$ (68). Los índices G2-LD y Q3 se emplearon para evaluar el supuesto de independencia local de los ítems(67). Para evaluar el ajuste de las personas (*Person Fit*) se empleó el estadístico Zh (Drasgow, Levine, y Williams)(71). El estadístico Zh se emplea para identificar a las personas cuya respuesta a los ítems no es consistente con el nivel de su rasgo latente. Un valor del Zh mayor de  $\pm 2$  refleja a las personas con un patrón de respuesta “atípico” o “inconsistente”(71).

El análisis estadístico se realizó en el programa R (R Core Team, 2017, Viena, Austria). Los paquetes empleados en el análisis de los datos fueron: *psych* para el análisis descriptivo y el AFE; *lavaan* y *semplot* para el AFC; *psych* y *semtools* para el análisis de fiabilidad; *mirt* para la calibración de los modelos de la TRI.

### **Aspectos éticos**

El estudio se realizó respetando los principios contenidos en los códigos de ética y las normas internacionales vigentes para las buenas prácticas en la investigación clínica y fue aprobado por el Comité de Ética del Instituto Mexicano del Seguro Social (número de

aprobación R-201-3201-085). Se obtuvo el consentimiento informado de todas las personas que participaron en el estudio.

## Resultados

### *Características generales de los participantes*

El estudio incluyó a 199 pacientes con OA de rodilla y 10 personas sin OA (personas sanas). Ciento dieciséis pacientes fueron reevaluados para la determinación de la estabilidad temporal (test-retest). La edad media de los participantes en la muestra de validación (n= 209) fue de 63 años (rango de edad de 34 a 90 años) y el 78.95% (n= 165) eran mujeres. La mediana de las puntuaciones del KOOS-12 en dolor, función, calidad de vida y resumen fue de 43.75, 37.5, 31.25 y 37.5, respectivamente. El porcentaje de items con valores perdidos fue bajo (<2%). Las características de los pacientes incluidos al inicio del estudio, en la muestra para el test-retest y la muestra de la sensibilidad al cambio se presentan en la tabla 1.

**Tabla 1. Características clínicas y demográficas de los participantes.**

	Muestra de validación (n= 199 pacientes y 10 sujetos sanos)	Muestra para el <i>test-retest</i> (n = 116 pacientes)	Muestra para la sensibilidad al cambio (n = 38 pacientes)
Mujeres, n(%)	165 (78.95)	100 (86.21)	32 (84.21)
Comorbilidad, n(%)			
Obesidad	108 (51.67)	70 (60.34)	19 (50)
Hipertensión	86 (41.15)	51 (43.97)	18 (47.37)
Diabetes mellitus	46 (22.01)	25 (21.55)	7 (18.42)
Dislipidemia	26 (12.44)	15 (12.93)	6 (15.79)
Osteoporosis	11 (5.29)	9 (7.75)	1 (2.63)
Hipotiroidismo	6 (2.87)	5 (4.31)	1 (2.63)
Años de educación	9 (6)	9 (6)	9 (9)
IMC	30.47 (7.23)	32.00 (6.40)	29.95 (8.69)
Edad	63.36 ± 11.83	62.26 ± 10.20	64.86 ± 10.21
ICOAP	52.27 (38.63)	57.13 ± 20.03	66.14 ± 21.75
KOOS-12			
Dolor	47.28 ± 24.47	43.75 ± 17.74	36.67 ± 23.97
Funcionamiento	37.5 (31.25)	36.11 ± 20.77	33.38 ± 25.88
Calidad de vida	31.25 (37.50)	28.25 ± 17.76	22.47 ± 18.25
Resumen	37.5 (29.16)	35.24 ± 16.33	30.84 ± 20.32
KOOS			
Dolor	49.68 ± 24.87	44.10 ± 18.55	39.54 ± 24.33
Síntomas	50 (39.28)	45.72 ± 19.48	42.19 ± 22.94
Actividades de la vida diaria	47.05 (35.29)	43.51 ± 20.66	39.31 ± 25.59
Actividades recreativas y deportivas	20 (27.73)	20 (30)	5 (25)
Valores perdidos			
Escala Dolor			
“Frecuencia del dolor”	0	1 (0.86%)	0
“Caminar”	0	0	0



“Subir/bajar escaleras”	1 (0.48%)	1 (0.86%)	1 (2.63%)
“Al sentarse o acostarse”	0	0	0
Escala funcionamiento			
“Levantarse de una silla”	1 (0.48%)	0	0
“Estar de pie”	2 (0.96%)	1 (0.86%)	0
“Entrar/salir de un auto”	2 (0.96%)	1 (0.86%)	0
“Girar/pivotear”	4 (1.91%)	4 (3.45%)	1 (2.63%)
Escala calidad de vida			
“Estar conciente del problema de su rodilla”	0	0	0
“Modificación de su estilo de vida”	4 (1.91%)	1 (0.86%)	0
“Falta de confianza en su rodilla”	1 (0.48%)	0	0
“Dificultad general de su rodilla”	1 (0.48%)	1 (0.86%)	0

---

Media  $\pm$  desviación estándar. Mediana (rango intercuartil)  
 IMC: Índice de Masa Corporal; ICOAP: Intermittent and Constant OsteoArthritis Pain scale; KOOS, Knee injury and Osteoarthritis Outcome Score.

### *Confiabilidad*

La tabla 2 muestra el análisis de confiabilidad de las subescalas del KOOS-12. El instrumento KOOS-12 mostró una alta consistencia interna. El alfa de Cronbach fue de 0.87, 0.91, 0.85 y 0.94 para las escalas de dolor, funcionamiento, calidad de vida y resumen, respectivamente. El coeficiente omega de McDonald de las escalas del KOOS-12 estuvieron comprendidos entre 0.85 (calidad de vida) y 0.90 (funcionamiento), lo cual indica una consistencia interna apropiada, es decir todos los ítems de cada una de las escalas miden el mismo constructo. El coeficiente de fiabilidad marginal de las tres escalas fue de 0.90. Lo cual demuestra que se obtienen mediciones precisas según este indicador de la TRI.

Ciento dieciséis personas fueron evaluadas en un lapso de 1 a 2 semanas de la evaluación inicial para evaluar la confiabilidad test-retest. La confiabilidad test-retest fue buena para la escala calidad de vida ( $ICC_{3,1} = 0.71$ ) y la medida de resumen ( $ICC_{3,1} = 0.71$ ). La confiabilidad test-retest de las escalas de funcionamiento ( $ICC_{3,1} = 0.60$ ) y dolor ( $ICC_{3,1} =$

0.63) fue aceptable. El error estándar de medida osciló entre 9.38 y 13.19. El cambio mínimo detectable para las escalas del KOOS-12 comprendió entre 28.32 y 36.56 (Tabla 2).

**Tabla 2. Análisis de la confiabilidad de los instrumentos KOOS y KOOS-12.**

	Consistencia interna (n= 209)			Estabilidad temporal (n =116)				
	Alfa de Cronbach (IC 95%)	Omega de McDonald	Coefficiente de correlación intraclase <sup>1</sup>	Error estándar de medida	Cambio mínimo detectable (individual)	Cambio mínimo detectable (grupal)	Promedio de las diferencias (Desviación estándar)	Limites de acuerdo del 95% (Bland-Altman)
<b>KOOS-12</b>								
Dolor	0.87 (0.84-0.90)	0.87	0.63 (0.51 to 0.73)	10.96	30.38	2.82	-0.28 (15.50)	-30.67; 30.09
Funcionamiento	0.91 (0.88-0.93)	0.90	0.60 (0.47 to 0.71)	13.19	36.56	3.39	1.45 (18.65)	-38.21; 35.25
Calidad de vida	0.85 (0.81-0.88)	0.85	0.71 (0.60 to 0.78)	10.22	28.32	2.63	0.66 (14.45)	-27.66; 28.99
<b>KOOS</b>								
Dolor	0.94 (0.93-0.95)		0.64 (0.53 to 0.74)	11.34	31.44	2.91	-0.44 (16.04)	-31.88; 30.99
Actividades de la vida diaria	0.98 (0.97-0.98)		0.63 (0.50 to 0.72)	12.78	35.42	3.28	0.43 (18.07)	-34.99; 35.86
Actividades deportivas y recreativas	0.95 (0.95-0.96)		0.64 (0.52 to 0.74)	11.98	33.22	3.09	0.19 (16.94)	-33.02; 33.41

<sup>1</sup> Two-way mixed-effect ANOVA model with interaction for the absolute agreement.  
IC 95%: Intervalo de confianza del 95%.

### *Validez basada en la estructura interna*

Debido a que cada una de las 3 escalas incluidas en el KOOS-12 se consideran escalas independientes, se realizó el análisis factorial confirmatorio en casa una de ellas. Todos los ítems se cargan fuertemente (cargas factoriales >0.7) en sus respectivos factores. El modelo de dolor KOOS-12 de un factor tuvo un ajuste adecuado del modelo (Figura 1). El AFC reveló que los modelos de un factor de las escalas de funcionamiento y calidad de vida del instrumento KOOS-12 mostraban un buen ajuste (Figuras 2 y 3).

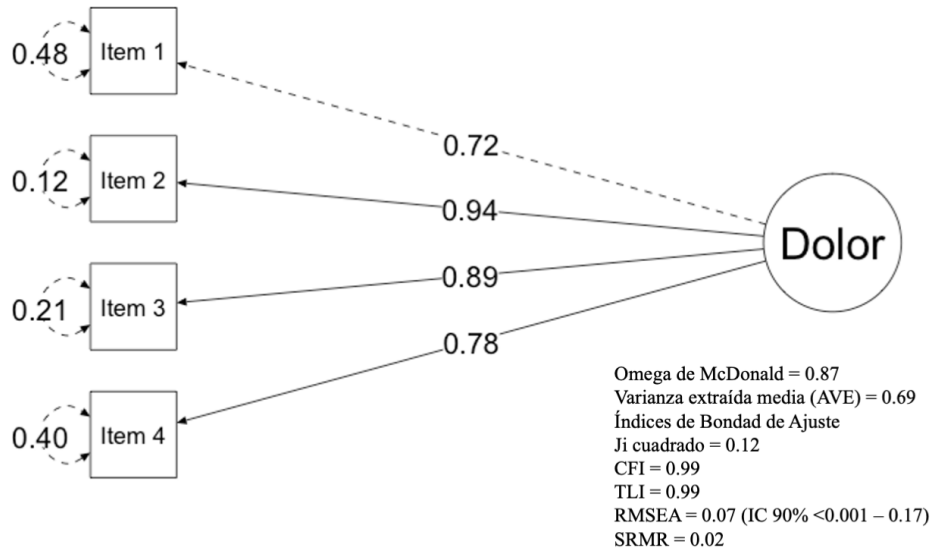


Figura 1. Representación (*Path diagram*) de la estructura interna y de los índices de bondad de ajuste de la escala de dolor del instrumento KOOS-12.

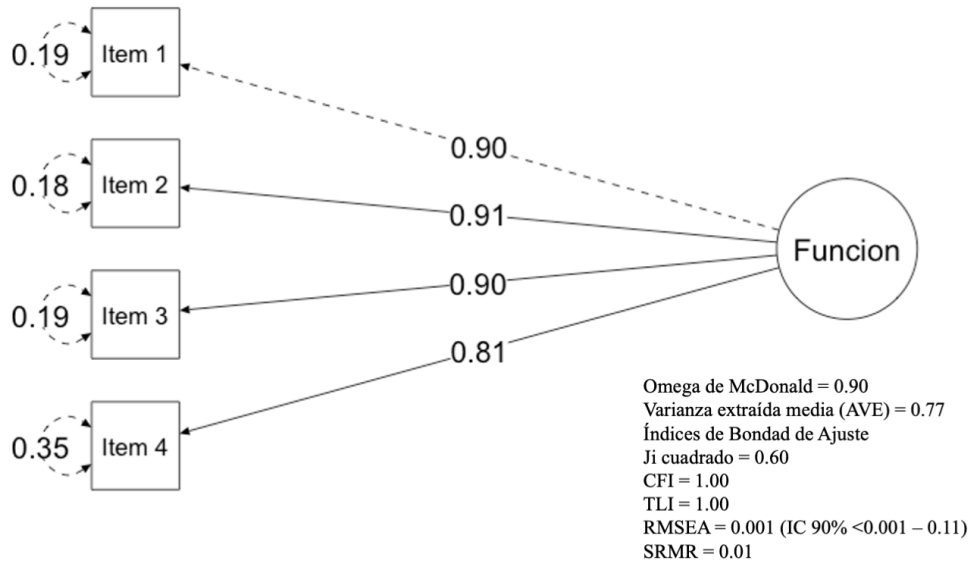


Figura 1. Representación (*Path diagram*) de la estructura interna y de los índices de bondad de ajuste de la escala de funcionamiento del instrumento KOOS-12.

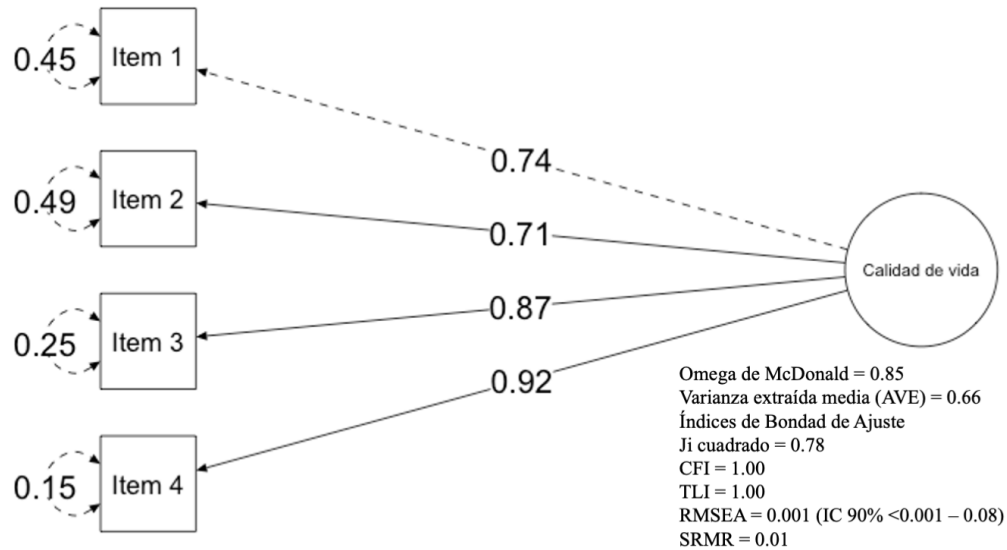


Figura 3. Representación (*Path diagram*) de la estructura interna y de los índices de bondad de ajuste de la escala de calidad de vida del instrumento KOOS-12.

### *Modelos de la Teoría de respuesta al ítem*

Los análisis de la TRI se realizaron por separado para las escalas de dolor, funcionamiento y calidad de vida. Las estimaciones de los parámetros obtenidos del modelo de crédito parcial generalizado (GPCM) y los estadísticos de ajuste de cada uno de los ítems se presentan en la Tabla 5. El rango de los parámetros de discriminación fue de 0.85 a 5.39, lo cual indica que los ítems permiten discriminar bien entre las personas con alta y baja afectación por la OA de rodilla. El ítem con más capacidad de distinguir entre pacientes fue el ítem 2 de la escala de dolor “dolor al caminar”. Por el contrario, el ítem 2 “modificaciones en el estilo de vida” fue el que demostró la menor discriminación. Todos los ítems incluidos en las tres escalas presentaron un buen ajuste a nivel de ítem. No se encontró la presencia de dependencia local entre los ítems. Las escalas de función y de calidad de vida tuvieron un buen ajuste general del modelo y el orden de los umbrales de las categorías para todos los ítems fue bueno

(figuras 4 y 5). Las estadísticas de ajuste de la persona detectaron <4% con patrones de respuesta "atípicos"(Tabla 3).

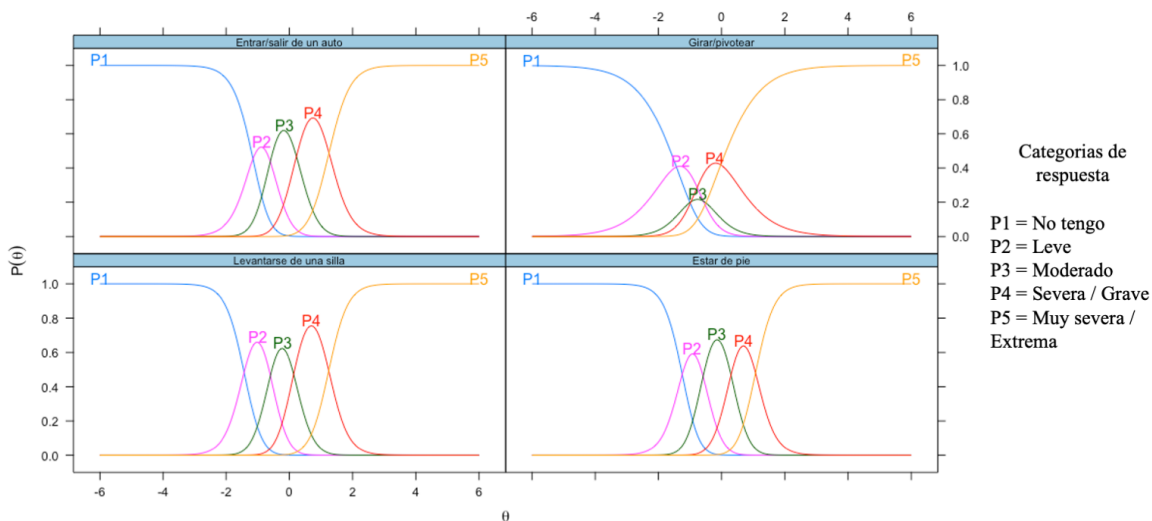


Figura 4. Curvas Características de las Categorías de respuesta de los ítems de la escala de funcionamiento. Se aprecia que en únicamente en el ítem "girar/pivotear" la tercera opción (P3) de respuesta no presentó una zona con mayor probabilidad de ser seleccionada

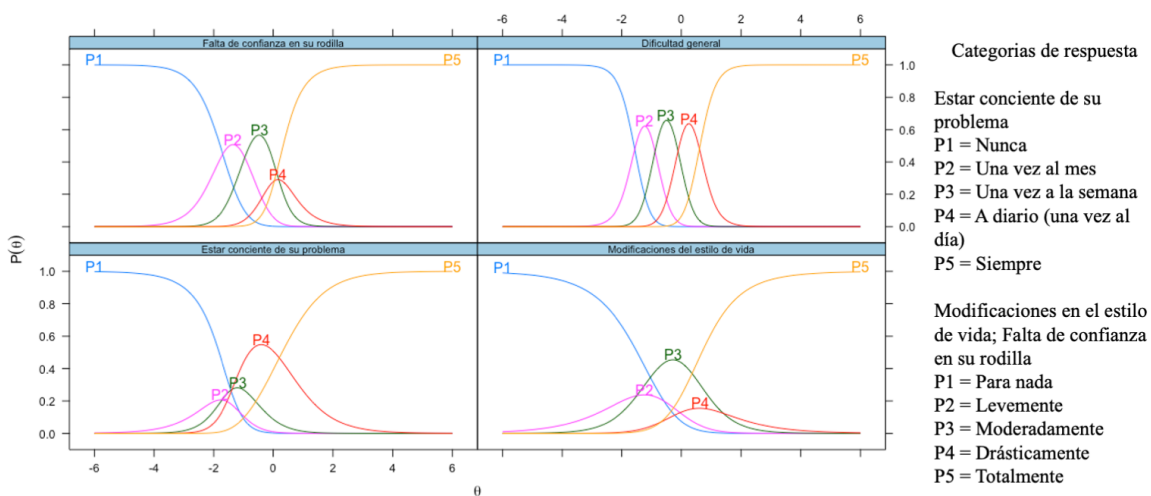


Figura 5. Curvas Características de las Categorías de respuesta de los ítems de la escala de calidad de vida. Se observa que en 3 de los 4 ítems, algunas opciones de respuesta no tienen una región en la escala logit con una mayor probabilidad de ser seleccionadas.

Tabla 3. Parámetros de discriminación y dificultad de los ítems, índices de bondad de ajuste de los modelos y coeficiente de confiabilidad marginal para cada una de las escalas del KOOS-12.												
Item	Discriminación del ítem ( $\alpha$ )	Dificultad del ítem (Thresholds) <sup>1</sup> .				Dificultad del ítem <sup>2</sup>	DIF	Item fit	Overall model fit	Person Fit		Coeficiente de confiabilidad marginal
		$\beta 1$ 2 vs. 1.	$\beta 2$ 3 vs. 2	$\beta 3$ 4 vs. 3	$\beta 4$ 5 vs. 4.					Misfitting responses <sup>4</sup>	Overfitting responses <sup>5</sup>	
<i>Dolor</i>												
1. "Frecuencia del dolor"	1.23	-1.37	-1.12	-1.31	0.86	-0.73	No	0.26	M2 = 0.01; RMSEA = 0.12 (0.04-0.21); SRMR = 0.03; TLI = 0.96; CFI = 0.98	1.9%	0%	0.90
2. "Caminar"	5.39	-0.92	-0.11	0.72	1.47	0.29	No	0.68				
3. "Subir/bajar escaleras"	2.71	-1.61	-0.89	-0.03	0.79	-0.43	No	0.49				
4. "Al sentarse o acostarse"	1.55	-0.69	0.01	1.10	1.69	0.53	No	0.16				
<i>Funcionamiento</i>												
1. "Levantarse de una silla"	3.20	-1.40	-0.61	0.11	1.22	-0.16	No	0.55	M2 = 0.30; RMSEA = 0.03 (<0.01-0.14); SRMR = 0.01; TLI = 0.99; CFI 0.99	3.9%	0%	0.90
2. "Estar de pie"	3.34	-1.25	-0.57	0.27	1.05	-0.12	No	0.57				
3. "Entrar/salir de un auto"	2.99	-1.17	-0.60	0.20	1.22	-0.08	No	0.57				
4. "Girar/pivotear"	1.44	-1.39	-0.60	-0.96	-0.04	-0.75	No	0.55				
<i>Calidad de vida</i>												
1. "Estar conciente del problema de su rodilla"	1.29	-1.12	-1.70	-1.35	0.20	-0.99	No	0.12	M2 = 0.56; RMSEA = <0.01 (<0.01-0.11); SRMR = 0.01; TLI = 1.00; CFI 1.00	1.9%	0%	0.90
2. "Modificación de su estilo de vida"	0.85	-0.57	-1.47	1.38	-0.68	-0.33	No	0.12				
3. "Falta de confianza en su rodilla"	2.04	-1.68	-0.94	0.17	0.05	-0.59	No	0.12				
4. "Dificultad general de su rodilla"	3.52	-1.55	-0.87	-0.10	0.59	-0.48	No	0.45				

<sup>1</sup> Cada ítem tiene cuatro umbrales que son: de  $\beta 1$  a  $\beta 4$ , donde  $\beta 1$  es el primer umbral (es decir, entre "Leve" y "Ninguno") y  $\beta 4$  es el cuarto umbral (es decir, entre las respuestas "Extremo" y "Severo").

<sup>2</sup> Las estimaciones de localización más altas indican ítems más difíciles.

<sup>3</sup> Corrección de Benjamini-Hochberg.

<sup>4</sup> Respuestas "misfit" = puntuación Zh inferior a -2.

<sup>5</sup> Respuestas "overfit" = puntuación Zh superior a 2.

DIF, funcionamiento diferencial de los ítems; RMSEA, error cuadrático medio de aproximación; SRMR, raíz cuadrada estandarizada; TLI, índice de Tucker Lewis; CFI, índice de ajuste comparativo.

La curva característica de los ítems es una función matemática que establece la relación existente entre las puntuaciones de las personas en la variable medida ( $\theta$ ) y la probabilidad de seleccionar esa opción de respuesta ( $P(\theta)$ ). Basándose en las curvas características de los ítems (Figura 6) se mostró que todos los ítems presentan umbrales ordenados y que únicamente la tercera categoría de respuesta “Una vez a la semana” no presentó en ninguna parte del recorrido en la escala de logit una mayor probabilidad de ser seleccionada. El ítem 1 "Frecuencia del dolor de rodilla" tuvo el parámetro de discriminación más bajo, lo que indica que este no discrimina tan bien entre los pacientes con baja y alta intensidad del dolor en comparación con los otros ítems. Los cuatro ítems de la escala de dolor abarcaban una amplia gama de dificultades que iban de -1.61 a 1.69. El ítem 4 "Dolor al sentarse o acostarse" fue el ítem con la dificultad más considerable de la escala de dolor, es decir, se requieren altos niveles de gravedad del dolor para que el paciente tenga una mayor probabilidad de seleccionar la última categoría de las opciones de respuesta "extrema". Por el contrario, el ítem con menor dificultad fue el ítem 3 "Dolor al subir o bajar escaleras", es decir, se requieren niveles muy bajos de gravedad del dolor de rodilla para que el paciente tenga una mayor probabilidad de seleccionar la categoría de respuesta "Ninguno". Por el contrario, la categoría 4 "extremo" tiene una alta probabilidad para los pacientes con mayor gravedad de la rodilla, pero la probabilidad disminuye a medida que lo hace la gravedad del dolor de rodilla. El ítem 2 "Dolor al caminar en plano" de la escala de dolor KOOS-12 proporcionó más información que los demás ítems. El ajuste global del modelo para la subescala de dolor fue aceptable ( $M2= 0.01$ ,  $RMSEA= 0.12$  y  $SRMR <0.05$ ; Tabla 3).

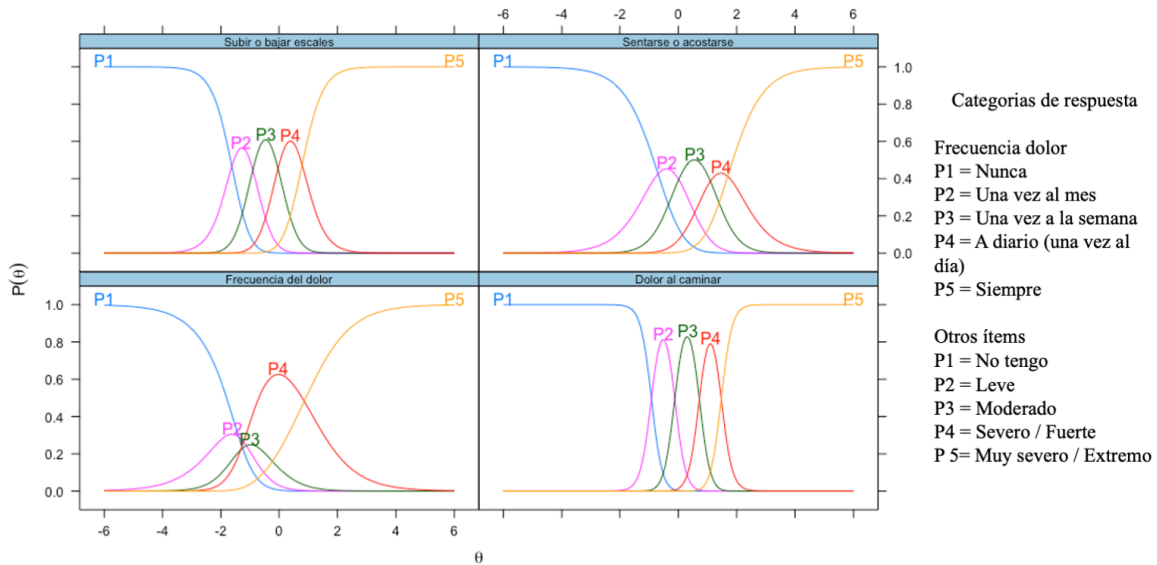


Figura 6. Curvas Características de las Categorías de respuesta de los ítems de la escala de calidad de vida. Se aprecia que la categoría P3 del ítem “frecuencia del dolor” no tiene una mayor zona de probabilidad de ser seleccionado en la escala logit.

Función de información de los ítems y de las escalas. La función de información permite evaluar la cantidad de información que puede proporcionarnos un test a cualquier nivel de la variable latente ( $\theta$ ) y en qué nivel de la variable latente la evaluación es más precisa. Cuanta más información tiene un ítem o un test, más precisa es la estimación de la habilidad para un sujeto. Es decir, cuanto mayor valor proporcione la función de información de un test para un nivel de la variable latente, mayor será la discriminación para ese nivel y menor su error de medida. En las figuras 7 a la 9 se presentan los gráficos de la función de información obtenido para los ítems y las escalas incluidas en el instrumento KOOS-12. Las estimaciones de  $\theta$  son más precisas en el rango de -2 a 1.5 para la escala de dolor, de -1.2 a 0 para la escala de funcionamiento y de -2 a 0 para la escala de calidad de vida. Como se puede observar, las escalas del KOOS-12 son más precisas para para niveles medio o bajos del dolor, funcionamiento y calidad de vida (figuras 7 a 9).

A nivel del ítem, las curvas de información indican para qué valor de  $\theta$  aporta más información el ítem y cuál es la magnitud de la información aportada para dicho valor de  $\theta$ .



Se puede observar que el “dolor al caminar” es el que proporciona el mayor nivel de información de la escala de dolor, seguido por el ítem “dolor al subir o bajar las escaleras”. Por el contrario, el ítem “dolor al sentarse o acostarse” contribuye pobremente y de forma aplanada a lo largo de todo el espectro del rasgo latente a la información que brinda el test. En el caso de la escala de funcionamiento, todos los ítems contribuyen a las estimaciones de  $\theta$ . En la escala de calidad de vida el ítem “dificultad general” es el que presenta la mayor información y el ítem “modificaciones al estilo de vida” es el ítem que presenta la menor información.

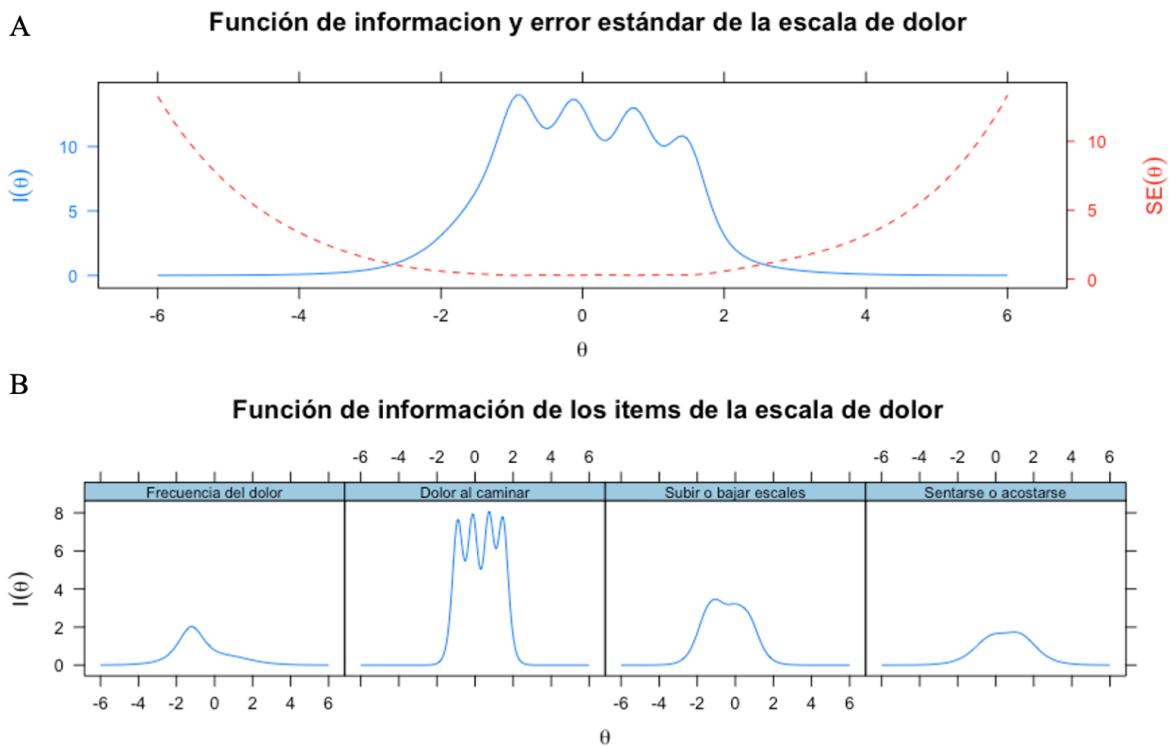


Figura 7A. Función de información del test para la escala de dolor del instrumento KOOS-12 y error típico de medida. 7B. Función de información para los ítems de la escala de dolor del instrumento KOOS-12.

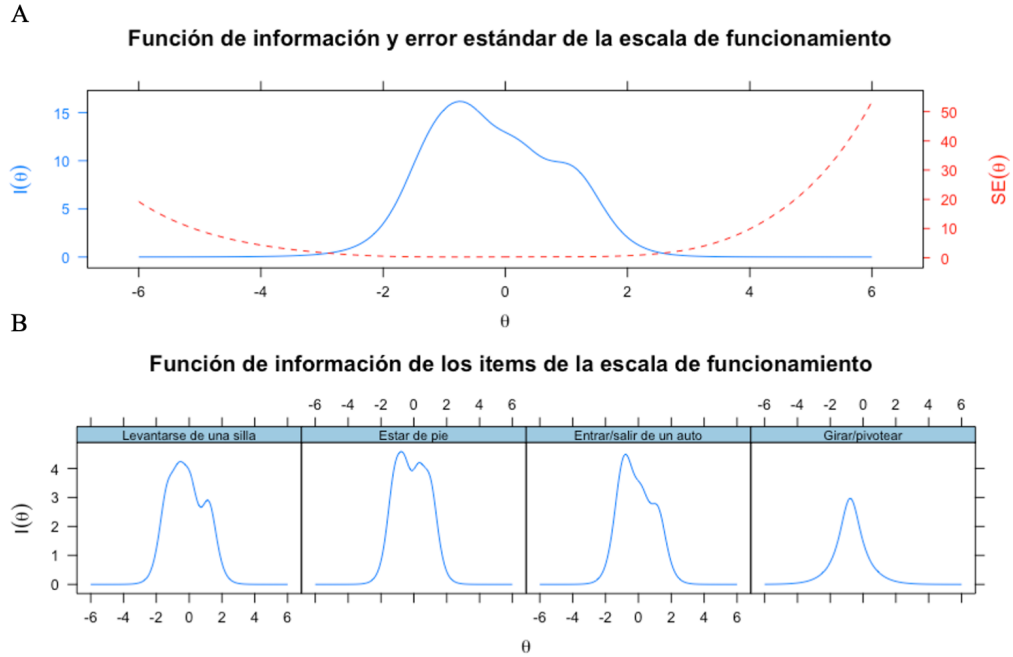


Figura 8A. Función de información del test para la escala de funcionamiento del instrumento KOOS-12 y error típico de medida. 8B. Función de información para los ítems de la escala de funcionamiento del instrumento KOOS-12.

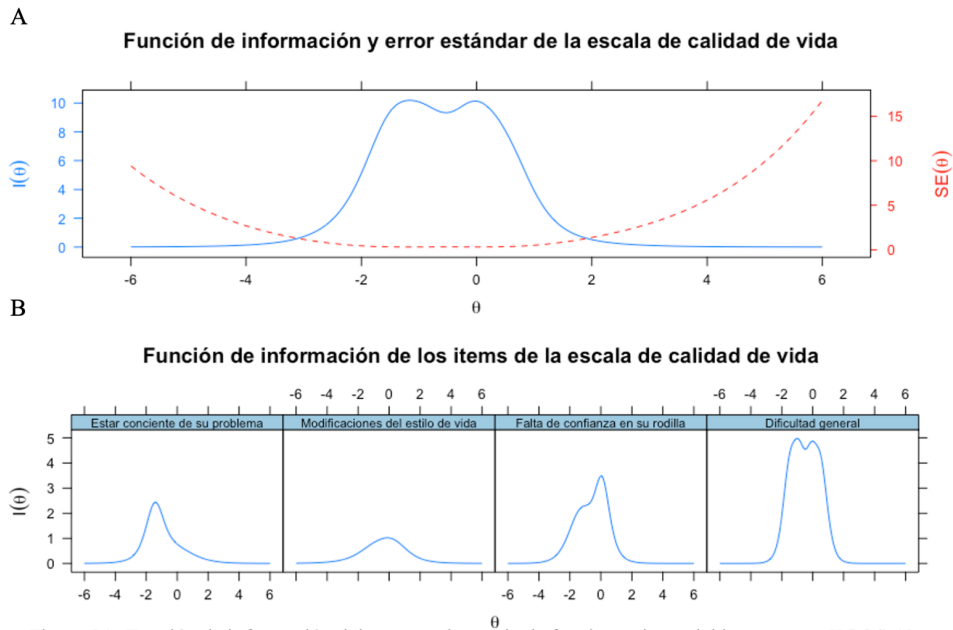


Figura 9A. Función de información del test para la escala de funcionamiento del instrumento KOOS-12 y error típico de medida 9B. Función de información para los ítems de la escala de funcionamiento del instrumento KOOS-12.

Parámetros de las personas. Los valores de la variable latente ( $\theta$ ) en las escalas de dolor, funcionamiento y calidad de vida se distribuyeron en valores logits de -3.89 a 1.29, -5.01 a 1.54 y -5.43 a 1.56, respectivamente. El porcentaje de personas presentaron un patrón de respuestas “atípicas” o “inconsistentes”, con un valor del estadístico  $Z_h < -3$ , fue del 1.9%, 3.9% y 1.9% para las escalas de dolor, funcionamiento y calidad de vida, respectivamente (tabla 3).

#### *Validez basada en la relación con otras variables*

Setenta y nueve por ciento de las hipótesis planteadas para la evaluación de la validez convergente se confirmaron. Similarmente, se confirmaron el 83% de las hipótesis para la evaluación de la validez divergente. La escala de dolor KOOS-12 mostró una correlación fuerte ( $\rho \geq 0.79$ ) con las escalas de dolor y actividades de la vida diaria del instrumento KOOS y con la escala de dolor ICOAP. La escala de funcionamiento del instrumento KOOS-12 mostró correlaciones fuertes ( $\rho \geq 0.80$ ) con las escalas de dolor, actividades de la vida diaria y actividades de recreativas y deportivas del instrumento KOOS y con la de dolor ICOAP. La escala calidad de vida del KOOS-12 presentó correlaciones fuertes con las escalas de actividades de la vida diaria y actividades recreativas y deportivas del instrumento KOOS y con la puntuación del instrumento IKDC. La puntuación de resumen del instrumento KOOS-12 se correlacionó fuertemente con la puntuación de las escalas de dolor, actividades de la vida diaria del KOOS y con la puntuación del instrumento IKDC y de la escala de dolor ICOAP. Las puntuaciones de las escalas del KOOS-12 y la puntuación del resumen del KOOS-12 entre la escala KOOS-12 presentaron una correlación moderada con las pruebas

basadas en el redimiento físico (coeficientes de correlación rho de Spearman entre 0,46 y 0,55; Tabla 4).

**Tabla 4. Validez del KOOS-12 basada en la relación con otras variables. Coeficiente de correlación de Spearman (Intervalo de confianza del 95%) entre el KOOS-12, instrumentos específicos para OA de rodilla y pruebas basadas en el funcionamiento físico.**

KOOS	Knee Injury and Osteoarthritis Outcome Score-12 item scale (KOOS-12).			
	Dolor	Función	Calidad de vida	Resumen
Dolor	0.94 (0.93 to 0.96)	0.83 (0.78 to 0.86)	0.66 (0.58 to 0.73)	0.89 (0.86 to 0.91)
Síntomas	0.73 (0.66 to 0.79)	0.78 (0.72 to 0.83)	0.66 (0.58 to 0.73)	0.79 (0.73 to 0.83)
Actividades de la vida diaria	0.85 (0.81 to 0.89)	0.94 (0.92 to 0.95)	0.73 (0.66 to 0.78)	0.93 (0.91 to 0.94)
Actividades recreativas y deportivas	0.61 (0.52 to 0.69)	0.80 (0.75 to 0.84)	0.70 (0.63 to 0.76)	0.77 (0.71 to 0.82)
Calidad de vida	0.63 (0.55 to 0.71)	0.73 (0.66 to 0.78)	1.00	0.87 (0.83 to 0.90)
IKDC	0.72 (0.65 to 0.78)	0.78 (0.72 to 0.83)	0.71 (0.63 to 0.77)	0.82 (0.77 to 0.86)
ICOAP	-0.79 (-0.83 to -0.73)	-0.73 (-0.79 to -0.66)	-0.67 (-0.73 to -0.58)	-0.80 (-0.84 to -0.74)
WHODAS 2.0				
Cognición	-0.20 (-0.30 to -0.05)	-0.32 (-0.45 to -0.19)	-0.22 (-0.35 to -0.08)	-0.28 (-0.40 to -0.14)
Mobilidad	-0.60 (-0.68 to -0.50)	-0.69 (-0.76 to -0.60)	-0.62 (-0.71 to -0.53)	-0.70 (-0.77 to -0.62)
Actividades de la vida diaria	-0.41 (-0.52 to -0.28)	-0.49 (-0.59 to -0.37)	-0.47 (-0.58 to -0.35)	-0.50 (-0.60 to -0.39)
Medidas basadas en el funcionamiento físico				
Test up and go	-0.46 (-0.58 to -0.33)	-0.55 (-0.65 to -0.42)	-0.46 (-0.58 to -0.32)	-0.54 (-0.65 to -0.42)
30-s chair test	0.50 (0.38 to 0.61)	0.51 (0.38 to 0.61)	0.49 (0.36 to 0.59)	0.54 (0.43 to 0.64)

KOOS, Knee Injury and Osteoarthritis Outcome Score; IKDC, International knee documentation committee (IKDC) subjective knee evaluation form; ICOAP Knee Intermittent and Constant Osteoarthritis Pain; WHODAS 2.0, World Health Organization Disability Assessment Schedule 2.0.

Se disponía de radiografías digitales de 172 pacientes para evaluar la validez de grupos conocidos. Como era esperado, los pacientes con mayor daño radiográfico basándose en la clasificación de Kellgren-Lawrence informaron de una mayor gravedad del dolor, un mayor deterioro en el funcionamiento y una peor calidad de vida. El análisis post-hoc demostró una disminución significativa de las escalas KOOS-12 y de la puntuación de resumen entre los sujetos con KOA frente a los controles sanos ( $p < 0.01$ ), grado 4 frente a grado 1 ( $p < 0.01$ ), grado 4 frente a grado 2 ( $p < 0.01$ ) y grado 4 frente a grado 3 ( $p < 0.05$ ). Por lo demás, no se encontraron diferencias significativas. El dolor, la función, la calidad de vida y la puntuación de resumen de la KOOS-12 fueron significativamente mayores en los pacientes no obesos que en los obesos. Los resultados de validez de los grupos conocidos fueron muy similares

utilizando las puntuaciones sumadas o los niveles Theta en la comparación de los grupos (tabla 5).

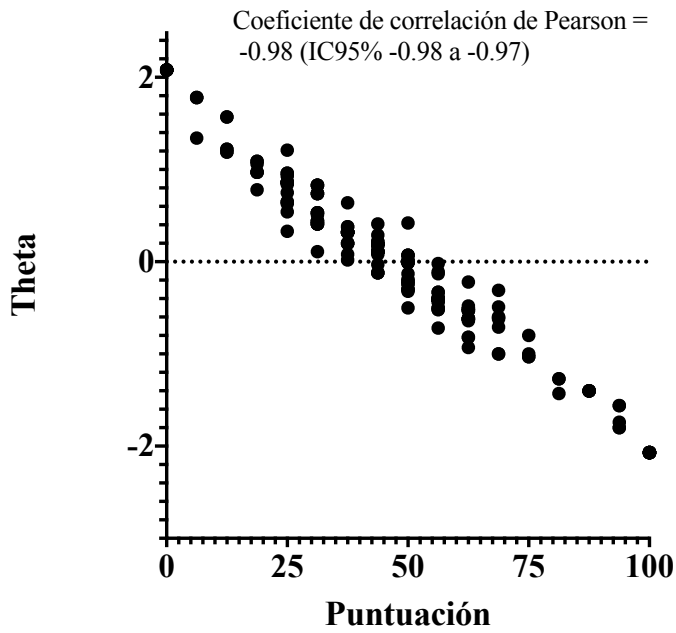
<b>Table 5. Resultados del análisis de la varianza (ANOVA) para las escalas del KOOS-12 y la puntuación de resumen empleando las puntuaciones sumadas-transformadas y la puntuación obtenida de la teoría de respuesta al ítem (TRI).</b>								
Clasificación de Kellgren-Lawrence	Grade 0 (n=10)	Grade 1 (n=21)	Grade 2 (n=70)	Grade 3 (n=47)	Grade 4 (n=24)	Valor de p	Eta cuadrada	
<b>KOOS-12 (score 0-100)</b>								
Pain	97.5±4.37	51.48±22.17	45.29±19.97	41.35±19.31	28.82±17.53	<0.001	0.37	
Function	96.62±6.62	47.91±28.25	42.05±22.06	34.70±22.26	21.09±16.77	<0.001	0.34	
Quality of life	88.12±12.99	44.94±24.97	32.67±22.32	26.68±15.86	10.41±11.60	<0.001	0.42	
Summary	93.75±6.80	48.11±22.24	40±19.56	34.24±16.31	19.44±12.67	<0.001	0.44	
<b>KOOS-12 (Theta or nivel de la variable latente)</b>								
Pain	-1.90 ± 0.25	-0.16 ± 0.84	0.05 ± 0.78	0.22 ± 0.75	0.79 ± 0.74	<0.001	0.36	
Function	-1.84 ± 0.34	-0.20 ± 1.04	-0.02 ± 0.71	0.24 ± 0.80	0.71 ± 0.76	<0.001	0.32	
Quality of life	-1.83 ± 0.43	-0.35 ± 0.80	0.04 ± 0.80	0.22 ± 0.65	0.88 ± 0.52	<0.001	0.40	
<b>Estado nutricional</b>		No obesos		Obesos		Valor de p	D de Cohen	
<b>KOOS-12 (score 0-100)</b>								
Pain		53.09 ± 25.28		41.33 ±21.58		<0.001	0.50	
Function		49.81 ± 26.39		34.37 ± 25.21		<0.001	0.59	
Quality of life		40.79 ± 26.94		28.41 ± 23.73		<0.001	0.49	
Summary		47.90 ± 24.04		34.70 ± 21.50		<0.001	0.57	
<b>KOOS-12 (Theta or latent trait level)</b>								
Pain		-0.22 ± 0.95		0.23 ± 0.85		<0.001	0.51	
Function		-0.27 ± 0.88		0.26 ± 0.92		<0.001	0.56	
Quality of life		-0.24 ± 0.91		0.22 ± 0.87		<0.001	0.52	
<b>Mean ± standard deviation</b>								

### Interpretación de los resultados

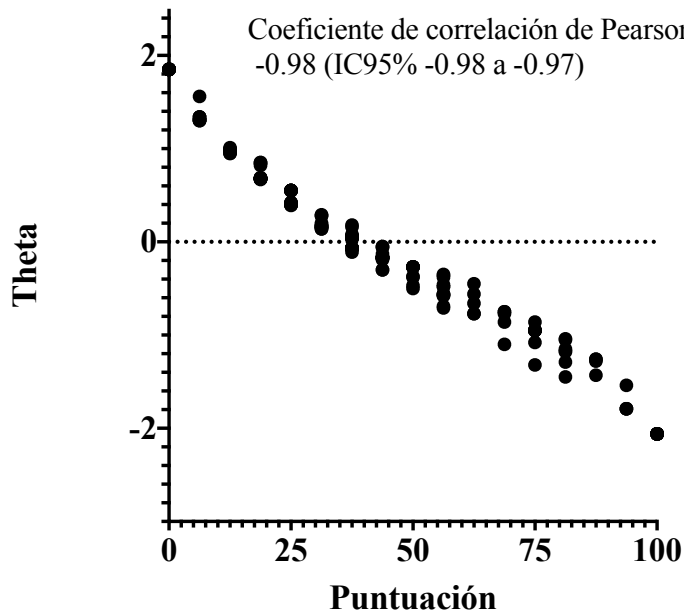
La puntuación de cada personas en cada una de las escalas se obtuvo por medio de 2 métodos:

1) empleando el método de suma y transformación a una escala de 0 a 100, 2) empleando el valor de Theta obtenido del modelo de respuesta de crédito parcial generalizado (la puntuación basada en la TRI). El coeficiente de correlación de Pearson entre las puntuaciones obtenidas por estos dos métodos fue de -0.983 (IC del 95%: 0.987 a -0.978), -0.982 (IC del 95%: -0.986 a -0.976) y -0.974 (IC del 95%: -0.980 a -0.966) para el dolor, la función y la calidad, respectivamente (figuras 10 a 12). Por lo tanto, es apropiado emplear la calificación obtenida por el método de transformación a 0 a 100.

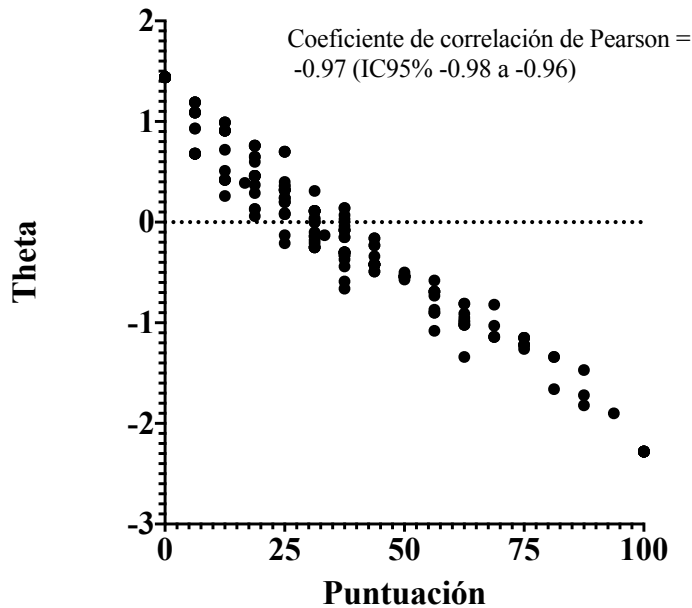
**Figura 10. Correlación entre la puntuación estimada y el valor Theta de la escala de dolor**



**Figura 11. Correlación entre la puntuación estimada y el valor Theta de la escala de Funcionamiento**



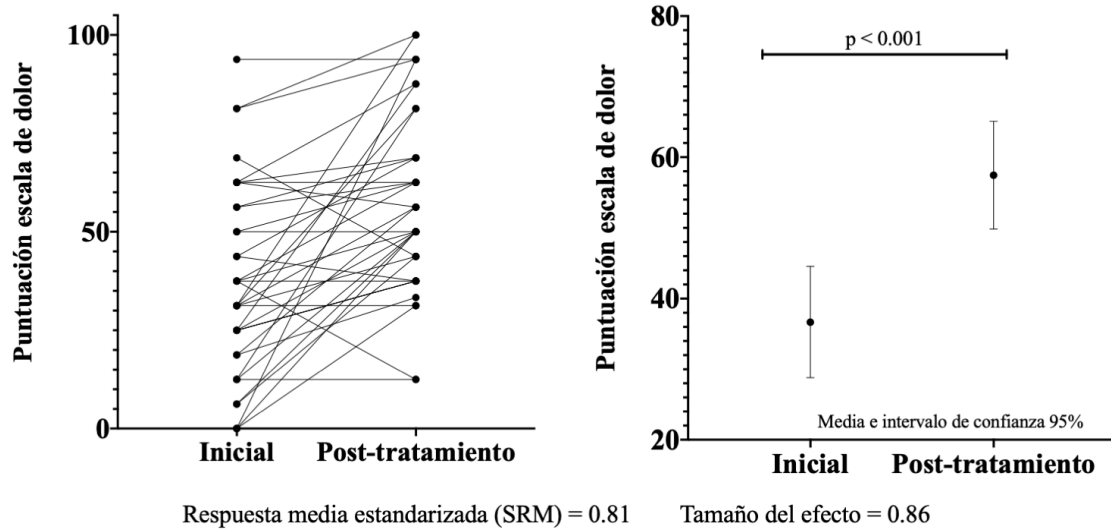
**Figura 12. Correlación entre la puntuación estimada y el valor Theta de la escala de Calidad de vida**



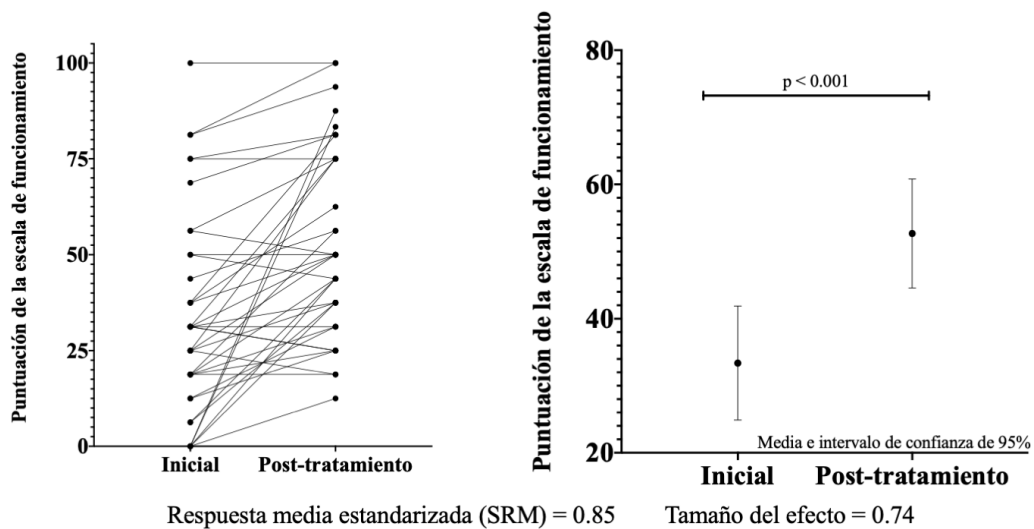
Efectos suelo y techo. Ninguna de las escalas del KOOS-12 presentó un efecto techo o suelo, es decir, que > 15% de los pacientes evaluados presenten la puntuación máxima o mínima, respectivamente. En los pacientes con OA de rodilla, el 2.01% (n= 4) presentaron la puntuación más alta (mejor resultado) en las escalas KOOS-12 y en la puntuación de resumen. Del mismo modo, el 2.01% (n= 4) presentaron la puntuación más baja (peor resultado) en las escalas KOOS-12 y en la puntuación de resumen.

Sensibilidad al cambio. La evaluación de la sensibilidad al cambio indicó que las escalas y la puntuación de resumen del KOOS-12 eran sensibles al cambio. Existió una mejoría significativa en las puntuaciones del KOOS-12 a las 8 semanas del tratamiento intra-articular (figuras 13 a 16).

**Figura 13. Cambio en la escala de dolor del KOOS-12 a las 8 semanas del tratamiento intra-articular**

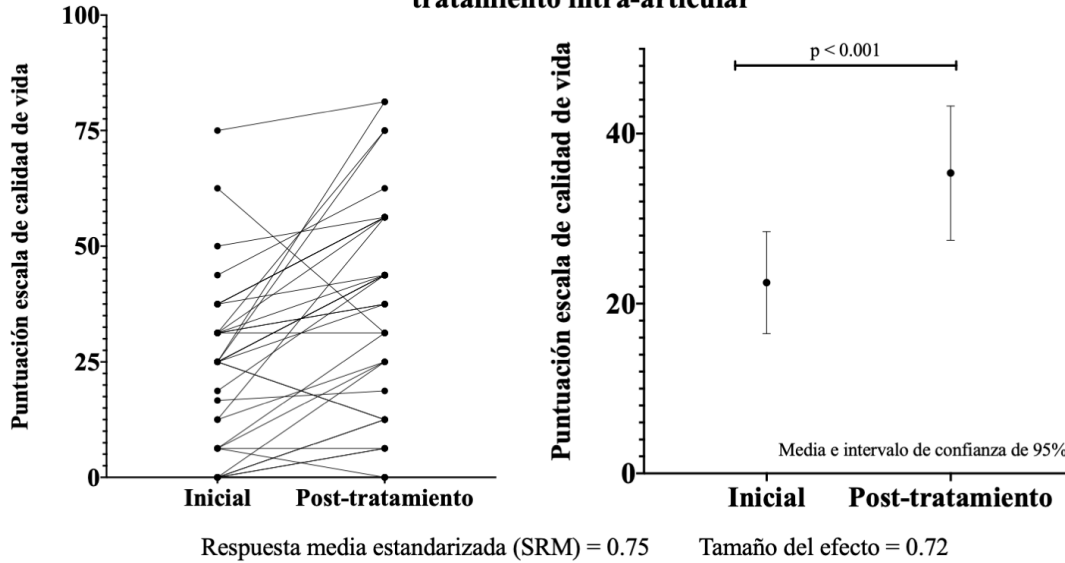


**Figura 14. Cambio en la escala de funcionamiento del KOOS-12 a las 8 semanas del tratamiento intra-articular**

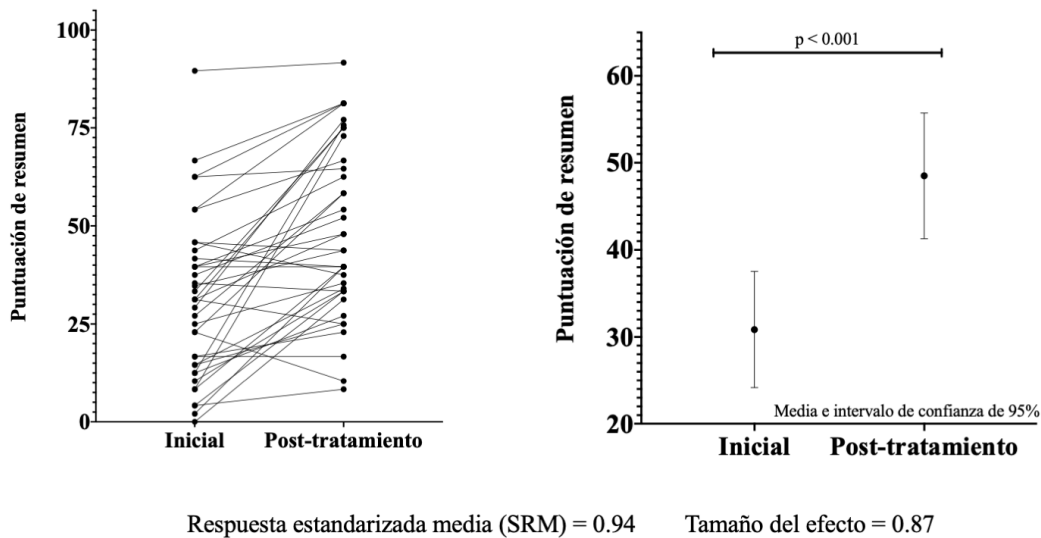




**Figura 15. Cambio en la escala de calidad de vida del KOOS-12 a las 8 semanas del tratamiento intra-articular**



**Figura 16. Cambio en la puntuación de resumen del KOOS-12 a las 8 semanas del tratamiento intra-articular**



La puntuación de resumen del instrumento KOOS-12 fue la que presentó el mayor tamaño del efecto de todas las escalas al tratamiento intra-articular (respuesta media estandarizada = 0.94, tamaño del efecto = 0.87). La puntuación de resumen de la diferencia del cambio a las 8 semanas del tratamiento intra-articular del instrumento KOOS-12 presentó una

correlación fuerte ( $r \geq 0.76$ ) y significativa con las escalas de dolor y actividades de la vida diaria del instrumento KOOS, la puntuación ICOAP y la puntuación IKDC. Por el contrario, no se demostró una correlación estadísticamente significativa con las pruebas basadas en el rendimiento físico (Tabla 6).

**Tabla 6. Sensibilidad del cambio de los instrumentos KOOS, KOOS-12, IKDC, ICOAP y las pruebas basadas en el rendimiento físico.**

	KOOS-12 Dolor	KOOS-12 Función	KOOS-12 Calidad de vida	KOOS-12 Resumen
KOOS Dolor	0.94 (0.89 to 0.97)	0.82 (0.68 to 0.90)	0.49 (0.21 to 0.70)	0.90 (0.83 to 0.95)
KOOS Actividades de la vida diaria	0.89 (0.79 to 0.94)	0.90 (0.82 to 0.95)	0.55 (0.28 to 0.74)	0.93 (0.87 to 0.96)
KOOS Calidad de vida	0.41 (0.11 to 0.65)	0.54 (0.27 to 0.73)	1	0.71 (0.50 to 0.84)
IKDC	0.74 (0.56 to 0.86)	0.69 (0.48 to 0.83)	0.49 (0.21 to 0.70)	0.76 (0.59 to 0.87)
ICOAP	-0.72 (-0.85 to -0.52)	-0.59 (-0.72 to -0.34)	-0.38 (-0.62 to -0.06)	-0.68 (-0.82 to -0.46)
TUG	-0.39 (-0.69 to 0.02)	-0.44 (-0.72 to -0.04)	-0.13 (-0.52 to 0.28)	-0.41 (-0.70 to 0.001)
30s CST	0.22 (-0.14 to 0.53)	0.18 (-0.18 to 0.50)	0.30 (-0.05 to 0.59)	0.27 (-0.09 to 0.57)

Timed up and go test, TUG; 30s CST, 30 second chair stan test.  
Coeficiente de correlación de Pearson (intervalo de confianza del 95%) entre las puntuaciones medias de cambio al inicio y dos meses después de la primera dosis del tratamiento intraarticular.

Funcionamiento diferencial del ítem (FDI). El análisis del funcionamiento diferencial del ítem evalúa si la probabilidad de seleccionar una respuesta por los pacientes con el mismo nivel del rasgo latente ( $\theta$ ) varía en función a alguna característica, por ejemplo, edad, sexo, nivel educativo o nivel socioeconómico. Los métodos para la detección del FDI comparan las respuestas dadas a un ítem por sujetos a los que se les estima un mismo nivel del rasgo latente en distintos. En el análisis del FID no se demostró que ninguno de los ítems de las escalas del instrumento KOOS-12 presenten un un funcionamiento diferencial del ítem por sexo, edad o nivel educativo. En la figura 17 se presenta el resultado gráfico del análisis del FDI de la escala de dolor.

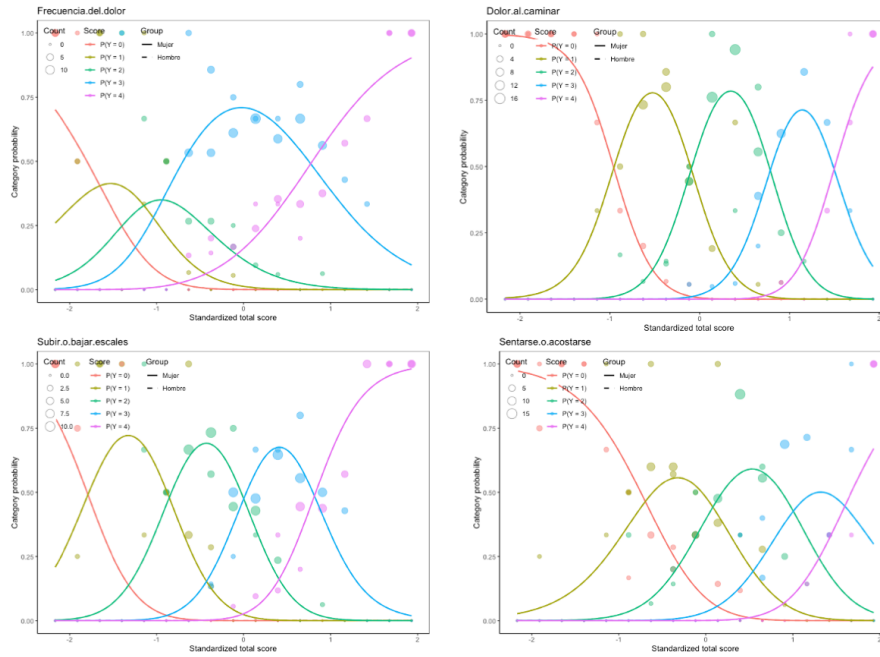


Figura 17. Representación gráfica del análisis del FDI de la escala del dolor del instrumento KOOS-12, se aprecia como existe una sobreposición de las curvas de hombre y mujeres lo cual demuestra la ausencia de FDI.

## Discusión

El instrumento KOOS-12 es una medida de desenlace desde la perspectiva del paciente, breve, que evalúa las opinión del pacientes sobre las dificultades que experimentan debido a los problemas con su rodilla y cubre los aspectos importantes de la OA de rodilla (dolor, limitaciones funcional y calidad de vida) (54). Actualmente, existen tres versiones en español del instrumento KOOS (adaptadas al español de España, Perú y Estados Unidos). En este estudio se evaluó las propiedades psicométricas de KOOS-12 utilizando la versión en español para Perú, en lugar de hacer otra traducción. No obstante, algunos pacientes necesitaron ayuda para aclarar algunas opciones de respuesta, por lo que se hicieron pequeñas modificaciones para mejorar la comprensión de las opciones de respuesta. Por ejemplo, algunos pacientes tuvieron dificultades para entender la diferencia entre "diariamente" y "siempre". Por lo tanto, se añadió la aclaración "Una vez al día" a la opción "diariamente". Del mismo modo, a las opciones de respuesta "severo" y "muy severo", se añadió la aclaración "Severo/Fuerte" y "Muy grave/Extremo".

Los resultados del estudio demuestran que la versión en español del instrumento KOOS-12 presenta un alto nivel de consistencia interna para evaluar los problemas de rodilla de los pacientes mexicanos con OA de rodilla. Esto confirma los hallazgos en un estudios previos que reportan un valor del coeficiente alfa de Cronbach en rangos de 0.75-0.82, 0.78-0.82, 0.80-0.84 y 0.90-0.93 para las escalas de dolor, función, calidad de vida y la puntuación de resumen, respectivamente(54). Los resultados de estabilidad temporal (test-retest) muestran un estabilidad temporal moderada para las escalas de dolor y función de la KOOS-12 (CCI

< 0.7) y buena para la escala de calidad de vida y la puntuación de resumen. A nuestro mejor conocimiento, ninguno de los estudios previos han informado de la confiabilidad test-retest del KOOS-12.

En el análisis basado en la TRI, únicamente la escala de dolor del instrumento KOOS-12 no presentó un buen ajuste global. La falta de ajuste global del modelo puede estar relacionado a la existencia de algunos ítems que no presentaron una zona de mayor probabilidad de ser seleccionados. Por ejemplo, la opción de respuesta "mensual" nunca fue la respuesta más probable para los pacientes en ningún punto del valor para el dolor. La explicación para este hallazgo, puede estar dada por la baja frecuencia en esta opción de respuesta del ítem 1. La frecuencia de la segunda opción de respuesta "mensual" fue del 7.18% y del 11.96% para la tercera opción de respuesta "semanal", que fueron mucho menores que el 46.89% de la cuarta opción de respuesta "diariamente". La baja frecuencia de estas categorías puede deberse al menor número de pacientes con OA de rodilla temprana. Por lo tanto, la evaluación de la escala de dolor podría ser adecuada con un mayor número de pacientes con OA de rodilla grado 1 de la clasificación de Kellgren-Lawrence.

La fuerte correlación entre la escala del dolor del instrumento KOOS-12 (compuesta de 4 ítems) y la escala del dolor del instrumento KOOS (compuesta de 9 ítems) indica que la varianza de la escala de dolor KOOS fue suficiente captada por los cuatro ítems de la escala de dolor KOOS-12. Estos resultados coinciden con los valores del coeficiente de correlación comunicados en dos estudios anteriores ( $r= 0.89-0.93$ ) (50, 51). Asimismo, se demostró una fuerte correlación ( $r= 0.94$ ) entre escala de funcionamiento del instrumento KOOS-12 y la escala de actividades de la vida diaria del instrumento KOOS, lo que coincide con los datos

comunicados anteriormente ( $r= 0.81-0.90$ )(50, 51). Por otro lado, se encontró una mayor correlación entre la escala de funcionamiento del instrumento KOOS-12 y la escala de actividades deportivas y recreativas del instrumento KOOS ( $r= 0,80$ ) en comparación con estudios anteriores ( $r= 0.61-0.71$ ) (50, 51). La validez discriminante del KOOS-12 se demostró por la baja correlación entre las escalas del instrumento KOOS-12 y las escalas cognitivas y de actividades de la vida diaria del instrumento WHODAS 2.0. Previamente, se demostró la baja correlación del instrumento KOOS-12 con la escala de salud mental del instrumento SF-36(54). La puntuación de resumen presentó evidencias basadas en la relación con otras variables que apoyan su validez de constructo. En este estudio, la puntuación de resumen del KOOS-12 se correlació fuertemente con las puntuaciones del IKDC, ICOAP y KOOS. Del mismo modo, Eckhar *et al.*(50) reportaron que la puntuación de resumen del KOOS-12 se correlacionó significativamente con la puntuación de los instrumentos KOOS, WOMAC y Oxford-12(50).

Dobson *et al.*(72) informaron de que las pruebas de bipedestación presentan la mejor evidencia, incluidas las pruebas TUG y el CST de 30 segundos, para la evaluación de la OA de rodilla. En este estudio, las escalas KOOS-12 mostraron una correlación moderada con las pruebas TUG y el CST de 30 segundos. Nuestros resultados son coherentes con los hallazgos reportados anteriormente, el TUG se correlaciona negativa y moderadamente con todas las escalas del instrumento KOOS ( $r= -0,66$  a  $-0,45$ ) y con el índice de Lequesne (73, 74) y presenta una correlación baja y positiva con el instrumento WOMAC ( $r < 0.3$ ) (74, 75). Hasta nuestro mejor conocimiento, ningún estudio previo evalúa la correlación entre el KOOS o KOOS-12 con la prueba CST de 30 segundos. Sin embargo, la información disponible sugiere una relación de baja a moderada entre las medidas autoinformadas y las

pruebas basadas en el rendimiento físico. Además, las medidas basadas en el rendimiento y las medidas de desenlace desde la perspectiva del paciente evalúan diferentes características de los pacientes.

Basándonos en la función de información de las escalas del KOOS-12, este instrumento es especialmente útil para evaluar los problemas asociados con la rodilla de leves a moderados. En afectación leve o moderada las escalas del KOOS-12 proporcionan una medición más precisa. La ausencia de funcionamiento diferencial del ítem por edad, sexo y nivel educativo permite su uso general en los pacientes referidos para su evaluación en la consulta externa de reumatología y ortopedia. La alta correlación entre la puntuación obtenida en una escala 0 a 100 y el valor  $\eta^2$  obtenido del modelo crédito parcial, demuestra que la escala de 0 a 100 es una medición confiable y válida del rasgo latente.

Las escalas del KOOS-12 fueron sensibles al cambio, con un tamaño del efecto entre moderados y grandes, a los 2 meses de la terapia intraarticular. Además, las escalas y la puntuación de resumen del instrumento KOOS-12 presentaron la misma sensibilidad al cambio el mismo que las escalas de dolor, síntomas y las actividades de la vida diaria del instrumento KOOS e IKDC, en términos del tamaño del efecto. A nuestro mejor conocimiento, éste es el primer estudio que examina la sensibilidad al cambio del KOOS-12 en pacientes con OA de rodilla tratados con terapia intraarticular. En los pacientes con reemplazo articular total de rodilla, el valor de la respuesta media estandarizada para las puntuaciones del KOOS-12 oscilaron entre 1.62-2.12 y la escala de calidad de vida fue la más sensible al cambio(54). En cambio, en los pacientes con tratamiento intraarticular, las escalas de dolor y función fueron las más sensibles al cambio. El presente estudio muestra

que la escala de dolor y la puntuación de resumen del instrumento KOOS-12 fueron más sensibles al cambio que las escalas del instrumento KOOS, en términos de la magnitud de la respuesta estandarizada media. De modo similar, la puntuación de resumen del instrumento KOOS-12 presentó un tamaño del efecto grande y una respuesta media estandarizada elevada posterior al reemplazo articular total de rodilla(50, 54).

Las ventajas que presenta el instrumento KOOS-12 para su empleo en la práctica clínica son: no se requiere de permiso para su uso, está disponible de forma gratuita, es fácil de entender y puntuar y el número de valores perdidos es bajo. En el ámbito clínico, la KOOS-12 es una medida de desenlace desde la perspectiva del paciente, breve pero exhaustiva, específica para los problemas de rodilla, con buenas propiedades psicométricas y proporciona una puntuación global del impacto en la rodilla, junto con medidas específicas de dominio



## **Conclusión**

La versión en español del HAQ-II presenta un grado adecuado de validez, confiabilidad y sensibilidad al cambio para la medición de los problemas de rodilla en pacientes mexicanos con osteoartritis de rodilla.

## Referencias

1. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis and rheumatism*. 2010;62(9):2569-81.
2. Kraus VB, Blanco FJ, Englund M, Karsdal MA, Lohmander LS. Call for standardized definitions of osteoarthritis and risk stratification for clinical trials and clinical use. *Osteoarthritis and cartilage*. 2015;23(8):1233-41.
3. Hu ZD. Red blood cell distribution width: a promising index for estimating activity of autoimmune disease. *J Lab Precis Med* 2016;1(4):1-6.
4. Reolid-Martínez R, Matos-Berroa S, Ayuso-Raya MC, de los Santos-Berrido E, Castro-Hurtado S, Arias-Alaminos M, et al. Capacidad funcional, características del dolor y tratamiento farmacológico en pacientes con artrosis de rodilla. *Rehabilitación* 2014;48(4):226-31.
5. Bernad-Pineda M, Heras-Sotos J, Garcés-Puentes MV. Calidad de vida en pacientes con artrosis de rodilla y/o cadera. *Rev Esp Cir Ortop Traumatol*. 2014;58(5):283-9.
6. Esquivel-Valerio JA, Orzua-de la Fuente WM, Vazquez-Fuentes BR, Garza-Elizondo MA, Negrete-Lopez R, Trevino-Montes DO, et al. The Impact of Osteoarthritis on the Functioning and Health Status of a Low-Income Population: An Example of a Disability Paradox. *Journal of clinical rheumatology : practical reports on rheumatic & musculoskeletal diseases*. 2017.
7. Burgos-Vargas R, Cardiel MH, Loyola-Sanchez A, De Abreu MM, Pons-Estel BA, Rossignol M, et al. Characterization of knee osteoarthritis in Latin America. A comparative analysis of clinical and health care utilization in Argentina, Brazil, and Mexico. *Reumatología clínica*. 2014;10(3):152-9.
8. De la Garza-Jiménez JL, Vázquez-Cruz E, Aguilar-Rivera T, Montiel-Jarquín A, Gutiérrez-Gabriel I, Barragán-Hervella R, et al. Calidad de vida en pacientes con limitación funcional de la rodilla por gonartrosis. En una unidad de primer nivel de atención médica. *Acta Ortopédica Mexicana*. 2013;27(6):367-70.
9. Espinosa R, Alcántar E, Álvarez E, Arce A, Díaz A, Esquivel J, et al. Clinical and epidemiologic characteristics of patients with osteoarthritis from the Mexican initiative of patients with osteoarthritis and rheumatoid arthritis (IMPACTAR). *Osteoarthritis and cartilage*. 2016;24:S219-S20.
10. Pelaez-Ballestas I, Sanin LH, Moreno-Montoya J, Alvarez-Nemegyei J, Burgos-Vargas R, Garza-Elizondo M, et al. Epidemiology of the rheumatic diseases in Mexico. A study of 5 regions based on the COPCORD methodology. *The Journal of rheumatology Supplement*. 2011;86:3-8.
11. Bennett K, Cardiel MH, Ferraz MB, Riedemann P, Goldsmith CH, Tugwell P. Community screening for rheumatic disorder: cross cultural adaptation and screening characteristics of the COPCORD Core Questionnaire in Brazil, Chile, and Mexico. The PANLAR-COPCORD Working Group. Pan American League of Associations for Rheumatology. Community Oriented Programme for the Control of Rheumatic Disease. *The Journal of rheumatology*. 1997;24(1):160-8.
12. Cardiel MH, Rojas-Serrano J. Community based study to estimate prevalence, burden of illness and help seeking behavior in rheumatic diseases in Mexico City. A COPCORD study. *Clinical and experimental rheumatology*. 2002;20(5):617-24.
13. Del Rio Najera D, Santana N, Pelaez-Ballestas I, Gonzalez-Chavez SA, Quinonez-Flores CM, Pacheco-Tena C. Prevalence of rheumatic diseases in Raramuri people in Chihuahua, Mexico: a community-based study. *Clinical rheumatology*. 2016;35 Suppl 1:43-52.
14. Loyola-Sanchez A, Richardson J, Pelaez-Ballestas I, Alvarez-Nemegyei J, Lavis JN, Wilson MG, et al. The impact of arthritis on the physical function of a rural Maya-Yucateco community and factors associated with its prevalence: a cross sectional, community-based study. *Clinical rheumatology*. 2016;35 Suppl 1:25-34.
15. Informe al ejecutivo federal y al congreso de la unión sobre la situación financiera y los riesgos del Instituto Mexicano del Seguro Social. 2012-2013. México: Instituto Mexicano del Seguro Social; 2013 [Primera edición: [Available from: <http://www.imss.gob.mx/sites/all/statics/pdf/informes/20122013/InformeCompleto.pdf>.
16. Informe al ejecutivo federal y al congreso de la unión sobre la situación financiera y los riesgos del Instituto Mexicano del Seguro Social. 2013-2014. México: Instituto Mexicano del Seguro Social; 2014 [Primera edición: [Available from: [http://www.imss.gob.mx/sites/all/statics/pdf/informes/20132014/21\\_InformeCompleto.pdf](http://www.imss.gob.mx/sites/all/statics/pdf/informes/20132014/21_InformeCompleto.pdf).

17. Informe al ejecutivo federal y al congreso de la unión sobre la situación financiera y los riesgos del Instituto Mexicano del Seguro Social. 2015-2016. México: Instituto Mexicano del Seguro Social; 2016 [Primera edición:[Available from: <http://www.imss.gob.mx/sites/all/statics/pdf/informes/20152016/21-InformeCompleto.pdf>.
18. Informe al ejecutivo federal y al congreso de la unión sobre la situación financiera y los riesgos del Instituto Mexicano del Seguro Social. 2014-2015. México: Instituto Mexicano del Seguro Social; 2015 [Primera edición:[Available from: <http://www.imss.gob.mx/sites/all/statics/pdf/informes/20142015/21-InformeCompleto.pdf>.
19. Macias-Hernandez SI. La discapacidad asociada a la osteoartritis. *Revista medica del Instituto Mexicano del Seguro Social*. 2014;52(5):484-5.
20. Holt HL, Katz JN, Reichmann WM, Gerlovin H, Wright EA, Hunter DJ, et al. Forecasting the burden of advanced knee osteoarthritis over a 10-year period in a cohort of 60-64 year-old US adults. *Osteoarthritis and cartilage*. 2011;19(1):44-50.
21. . !!! INVALID CITATION !!!
22. Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics. *Osteoarthritis and cartilage*. 2004;12 Suppl A:S39-44.
23. Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis and cartilage*. 2010;18(1):24-33.
24. Montoya-Verdugo CA. Evolución clínica de los pacientes con gonartrosis tratados mediante la aplicación de colágeno polivinilpirrolidona. *Ortho-tips*. 2012;8(2):96-102.
25. Corti MC, Rigon C. Epidemiology of osteoarthritis: prevalence, risk factors and functional impact. *Aging clinical and experimental research*. 2003;15(5):359-63.
26. Perez Martin A. Clínica. Localizaciones: rodilla, cadera, manos, columna, otras localizaciones. *Atención primaria*. 2014;46 Suppl 1:11-7.
27. Castano Carou A, Pita Fernandez S, Pertega Diaz S, de Toro Santos FJ, Grupo de estudio E. Clinical profile, level of affection and therapeutic management of patients with osteoarthritis in primary care: The Spanish multicenter study EVALUA. *Reumatología clínica*. 2015;11(6):353-60.
28. Feinstein AR. Clinimetric perspectives. *Journal of chronic diseases*. 1987;40(6):635-40.
29. Lara-Muñoz MdC, Ortega-Soto H. ¿La clinimetría o la psicometría? *Medición en la práctica psiquiátrica*. *Salud Men*. 1995;18(4):33-40.
30. Lifshitz A. *La nueva clínica*. México: Intersistemas, S.A. de C.V.; 2014.
31. Fava GA, Ruini C, Rafanelli C. Psychometric theory is an obstacle to the progress of clinical research. *Psychother Psychosom*. 2004;73(3):145-8.
32. Wright JG, Feinstein AR. A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales. *J Clin Epidemiol*. 1992;45(11):1201-18.
33. Feinstein AR. *Clinimetrics*. New Haven: Yale University Press; 1985.
34. Streiner DL. Clinimetrics vs. psychometrics: an unnecessary distinction. *J Clin Epidemiol*. 2003;56(12):1142-5; discussion 6-9.
35. de Vet H, Terwee CB, Bouter LM. Clinimetrics and psychometrics: two sides of the same coin. *J Clin Epidemiol*. 2003;56:1146-7.
36. Meneses J, Barrios M, Albert B, Coscolluela A, Lozano LM, Turbany J, et al. *Psicometria*. Barcelona: Editorial UOC; 2013.
37. Eiliana M-R. Escalas o índices para la medición de constructos: el dilema del analista de datos. *Avances en medición*. 2008;6:15-24.
38. Alonso-Caballero J, Ferrer-Fores M. *Resultados reportados por los pacientes (PROs)*. Madrid, España: Ergon; 2017.
39. Horta-Baas G. Validación de una versión en español del Health Assessment Questionnaire- II para la evaluación de la capacidad funcional en pacientes mexicanos con artritis reumatoide. *Reumatología clínica*. 2020 <https://doi.org/10.1016/j.reuma.2020.11.005>.
40. Espinosa JC, Herrera-Rojas AN. Evaluación psicométrica de un instrumento construido a través de enlace empírico (MMPI) *Avances en medición*. 1:91-104.
41. Castellvi P, Ferrer M, Alonso J, en nombre del Comité Científico de Biblio PRO. [The patient-reported outcomes in research: definition, impact, classification, measurement and assessment]. *Medicina clínica*. 2013;141(8):358-65.
42. Pelaez-Ballestas I. Medidas de desenlace reportadas por el paciente: ¿cuál es su importancia? *Reumatología clínica*. 2012;8(3):105-6.

43. Gomez-Valero S, Garcia-Perez F, Florez-Garcia MT, Miangolarra-Page JC. Revision sistematica de los cuestionarios autocumplimentados adaptados al espanol para la valoracion funcional de pacientes con afecciones de rodilla. *Rev Esp Cir Ortop Traumatol*. 2017;61(2):96-103.
44. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. *J Orthop Sports Phys Ther*. 1998;28(2):88-96.
45. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. *Health Qual Life Outcomes*. 2003;1:64.
46. Vázquez-Barquero JL, Herrera-Castanedo H, Vázquez-Bourgón E, Gaité-Pindado L. Cuestionario para la Evaluación de Discapacidad de la Organización Mundial de la Salud (World Health Organization Disability Assessment Schedule II).
47. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes*. 2003;1:17.
48. Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS)--validation of a Swedish version. *Scand J Med Sci Sports*. 1998;8(6):439-48.
49. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Annals of the rheumatic diseases*. 1957;16(4):494-502.
50. Eckhard L, Munir S, Wood D, Talbot S, Brighton R, Walter B, et al. The KOOS-12 shortform shows no ceiling effect, good responsiveness and construct validity compared to standard outcome measures after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc*. 2020.
51. Gandek B, Roos EM, Franklin PD, Ware JE, Jr. Item selection for 12-item short forms of the Knee injury and Osteoarthritis Outcome Score (KOOS-12) and Hip disability and Osteoarthritis Outcome Score (HOOS-12). *Osteoarthritis Cartilage*. 2019;27(5):746-53.
52. Lassere MN. Imaging: the need for standardization. *Best Pract Res Clin Rheumatol*. 2008;22(6):1001-18.
53. Collins NJ, Roos EM. Patient-reported outcomes for total hip and knee arthroplasty: commonly used instruments and attributes of a "good" measure. *Clin Geriatr Med*. 2012;28(3):367-94.
54. Gandek B, Roos EM, Franklin PD, Ware JE, Jr. A 12-item short form of the Knee injury and Osteoarthritis Outcome Score (KOOS-12): tests of reliability, validity and responsiveness. *Osteoarthritis Cartilage*. 2019;27(5):762-70.
55. Alageel M, Al Turki A, Alhandi A, Alohalı R, Alsalem R, Aleissa S. Cross-Cultural Adaptation and Validation of the Arabic Version of the Intermittent and Constant Osteoarthritis Pain Questionnaire. *Sports Med Int Open*. 2020;4(1):E8-E12.
56. Panah SH, Baharlouie H, Rezaeian ZS, Hawker G. Cross-cultural adaptation and validation of the Persian version of the Intermittent and Constant Osteoarthritis Pain Measure for the knee. *Iran J Nurs Midwifery Res*. 2016;21(4):417-23.
57. Baidya O, Prabhakar R, Wadhwa M, Baidya P. Cross-cultural translation and validation of the Hindi version of the intermittent and constant osteoarthritis pain scale in knee osteoarthritis patients. *Ortho & Rheum Open Access*. 2018;10:1-5.
58. Mokkink LB, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, de Vet HC, et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments 2019 [Available from: [https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist\\_final.pdf](https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf)].
59. Reeve BR, Fayers P. Applying item response theory modelling for evaluating questionnaire itemscale properties. In: Fayers P, Hays H, editors. *Assessing Quality of Life in Clinical Trials: Methods and Practice*. 2nd ed. New York: Oxford University Press; 2005:55-73.
60. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum*. 1986;29(8):1039-49.
61. Vaquero J, Longo UG, Forriol F, Martinelli N, Vethencourt R, Denaro V. Reliability, validity and responsiveness of the Spanish version of the Knee Injury and Osteoarthritis Outcome Score (KOOS) in patients with chondral lesion of the knee. *Knee Surg Sports Traumatol Arthrosc*. 2014;22(1):104-8.
62. Bond M, Davis A, Lohmander S, Hawker G. Responsiveness of the OARSI-OMERACT osteoarthritis pain and function measures. *Osteoarthritis Cartilage*. 2012;20(6):541-7.
63. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis

- Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS). *Arthritis care & research*. 2011;63 Suppl 11:S208-28.
64. Vazquez-Barquero JL, Vazquez Bourgon E, Herrera Castanedo S, Saiz J, Uriarte M, Morales F, et al. Version en lengua española de un nuevo cuestionario de evaluación de discapacidades de la OMS (WHO-DAS-II): fase inicial de desarrollo y estudio piloto. Grupo Cantabria en Discapacidades. *Actas Esp Psiquiatr*. 2000;28(2):77-87.
65. Dobson F, Hinman RS, Roos EM, Abbott JH, Stratford P, Davis AM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. *Osteoarthritis and cartilage*. 2013;21(8):1042-52.
66. Steichen TJ, Cox NJ. Concordance Correlation coefficient. *Stata J*. 1998;43:35-9.
67. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*. 2007;45(5 Suppl 1):S22-31.
68. Horta-Baas G. Reliability and Validity of the Spanish Version of the Medical Outcomes Study Pain Severity Scale in Mexican Patient with Rheumatic Diseases. *J Clin Rheumatol*. 2020.
69. Dinno A. Nonparametric pairwise multiple comparisons in independent groups using Dunn's test. *Stata J*. 2015;15(1):292-300.
70. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
71. Felt JM, Castaneda R, Tiemensma J, Depaoli S. Using Person Fit Statistics to Detect Outliers in Survey Research. *Front Psychol*. 2017;8:863.
72. Dobson F, Hinman RS, Hall M, Terwee CB, Roos EM, Bennell KL. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis: a systematic review. *Osteoarthritis Cartilage*. 2012;20(12):1548-62.
73. Marconcin P, Espanha M, Yázigí F, Teles J. Predictor of Timed "Up-and-Go" Test in Elderly with Knee Osteoarthritis. In: Cabri J, Pezarat-Correia P, editors. *Proceedings of the 3rd International Congress on Sport Sciences Research and Technology Support*. 1. Lisbon, Portugal: Science and Technology Publications, Lda; 2015.
74. de Rezende MU, de Farias FES, da Silva CAC, Cernigoy CHA, de Camargo OP. Objective functional results in patients with knee osteoarthritis submitted to a 2-day educational programme: a prospective randomised clinical trial. *BMJ Open Sport Exerc Med*. 2016;2(1):e000200.
75. Gandhi R, Tsvetkov D, Davey JR, Syed KA, Mahomed NN. Relationship between self-reported and performance-based tests in a hip and knee joint replacement population. *Clin Rheumatol*. 2009;28(3):253-7.

## **Anexos**

## Anexo 1. Encuesta KOOS-12 de evaluación de la rodilla

**INSTRUCCIONES:** Esta es una encuesta que recoge su opinión sobre el estudio de su rodilla. La información que nos proporcione, servirá para saber como se siente con respecto a su rodilla y qué tan capaz es de realizar sus actividades diarias. Responda a cada pregunta marcando la casilla apropiada, sólo una casilla por pregunta. Si tiene alguna duda sobre como responder alguna pregunta, por favor proporcione la mejor respuesta posible.

### Dolor

D1. ¿Qué tan seguido siente dolor en la rodilla?

Nunca	Una vez al mes	Una vez a la semana	A diario (Una vez al día)	Siempre
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

¿Cuánto dolor de rodilla ha experimentado usted en la última semana (los 7 días previos) al realizar las siguientes actividades?

D2. Al caminar en una superficie plana

No tengo	Leve	Moderado	Severo/Fuerte	Muy severo/Extremo
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

D3. Al subir o bajar las escaleras

No tengo	Leve	Moderado	Severo/Fuerte	Muy severo/Extremo
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

D4. Al estar sentado o recostado

No tengo	Leve	Moderado	Severo/Fuerte	Muy severo/Extremo
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

### Actividades diarias

Las siguientes preguntas indagan sobre sus actividades físicas diarias. Es decir, su capacidad de moverse y valerse por sí mismo. Para cada una de las actividades mencionadas a continuación, indique el grado de dificultad experimentado en la última semana (los 7 días previos) con respecto a su rodilla.

A1. Al levantarse después de estar sentado

No tengo	Leve	Moderada	Severa/Grave	Muy severa/Extrema
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

A2. Al estar de pie

No tengo	Leve	Moderada	Severa/Grave	Muy severa/Extrema
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

A3. Al subir o bajar de un carro (auto)

No tengo	Leve	Moderada	Severa/Grave	Muy severa/Extrema
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

A4. Girar/impulsarse sobre la rodilla afectada

No tengo	Leve	Moderada	Severa/Grave	Muy severa/Extrema
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

### Calidad de vida

Q1. ¿Qué tan seguido es consciente del problema de su rodilla?

Nunca	Una vez al mes	Una vez a la semana	A diario (Una vez al día)	Siempre
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

Q2. ¿Ha modificado su estilo de vida para evitar actividades que podrían dañar su rodilla?

Para nada	Levemente	Moderadamente	Drásticamente	Totalmente
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

Q3. ¿Qué tan preocupado está usted con la falta de seguridad de su rodilla?

Para nada	Levemente	Moderadamente	Drásticamente	Totalmente
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

Q4. En general: ¿Cuánta dificultad tiene con su rodilla?

No tengo	Leve	Moderado	Severo	Muy severo
<input type="checkbox"/> <sup>0</sup>	<input type="checkbox"/> <sup>1</sup>	<input type="checkbox"/> <sup>2</sup>	<input type="checkbox"/> <sup>3</sup>	<input type="checkbox"/> <sup>4</sup>

**Muchas gracias por contestar todas las preguntas de este cuestionario**



## Puntuación de la escala KOOS-12

Cada uno de las subescalas del instrumento es calificada en una escala de 0 a 100, en donde una puntuación mayor indica una mejor funcionalidad de la rodilla. El cálculo establecido en el manual de usuario del cuestionario puede realizar con base en una plantilla de Excel o por medio de un cálculo manual de acuerdo a las siguientes fórmulas:

	Fórmula
Dolor	$100 - \text{media}(D1:D4) / 4 * 100$
Funcionamiento	$100 - \text{media}(A1:A4) / 4 * 100$
Calidad de vida	$100 - \text{media}(Q1:Q4) / 4 * 100$
Resumen	Media(dolor, funcionamiento, calidad de vida)

## Anexo 2. Carta de aprobación del comité local de ética

27/12/2018

SIRELCIS



INSTITUTO MEXICANO DEL SEGURO SOCIAL  
DIRECCIÓN DE PRESTACIONES MÉDICAS



### Dictamen de Autorizado

Comité Local de Investigación en Salud 3201 con número de registro 17 CI 31 050 062 ante COFEPRIS y número de registro ante CONBIOÉTICA CONBIOÉTICA 31 CEI 002 2018073.  
H GRAL REGIONAL NUM 1

FECHA Jueves, 27 de diciembre de 2018.

**M.C. GABRIEL JOSE HORTA BAAS**  
**PRESENTE**

Tengo el agrado de notificarle, que el protocolo de investigación con título:

**Evaluación de las propiedades psicométricas de la versión en español del International Knee Documentation Committee subjective knee evaluation form (IKDC) y del Knee Injury and Osteoarthritis Outcome score (KOOS) para la evaluación de los sujetos con osteoartritis primaria de la rodilla**

que sometió a consideración para evaluación de este Comité Local de Investigación en Salud, de acuerdo con las recomendaciones de sus integrantes y de los revisores, cumple con la calidad metodológica y los requerimientos de ética y de investigación, por lo que el dictamen es **A U T O R I Z A D O**, con el número de registro institucional:

No. de Registro  
R-2018-3201-085

ATENTAMENTE

**DR. JACINTO HERRERA LEÓN**  
Presidente del Comité Local de Investigación en Salud No. 3201