

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE MEDICINA



TESIS DOCTORAL

**Caracterización de la calidad de vida relacionada con la salud
en pacientes con artrosis de miembro inferior y validación de
nuevas herramientas para su elevación**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Roberto Carlos García Maroto

Directores

Jesús Martín Fernández
Fernando Marco Martínez

Madrid

© Roberto Carlos García Maroto, 2022

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE MEDICINA



TESIS DOCTORAL

Caracterización de la Calidad de Vida Relacionada con la Salud
en pacientes con artrosis de miembro inferior y validación de
nuevas herramientas para su evaluación

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Roberto Carlos García Maroto

DIRECTORES

Dr. Jesús Martín Fernández
Prof. Fernando Marco Martínez

AGRADECIMIENTOS

A la hora de emprender un trabajo ambicioso como es la presentación y defensa de una tesis doctoral es preciso un impulso e inquietud personal, pero obligatoriamente este debe ser respaldado, apoyado e incluso diría influenciado por la personas que intervienen en nuestra formación y desarrollo profesional. En este apartado quiero expresar un sincero agradecimiento al Profesor Fernando Marco Martínez, por su impulso continuado a lo largo de mi formación profesional, para motivar la faceta investigadora como un elemento fundamental de nuestra profesión.

El desarrollo de esta tesis no hubiera sido posible sin la ayuda del Dr. Jesús Martín Fernández, que fue la persona que me acerco a un grupo de trabajo con el que colaborar para realizar las publicaciones previas que justifican el proyecto. Todos los datos del trabajo de campo fueron recogidos gracias al esfuerzo de los investigadores y colaboradores de los proyectos de investigación PI PI1300560, PI1300518 y PI1300648 (financiados por el Instituto de Salud Carlos III y el Fondo Europeo de Desarrollo Regional (FEDER), “Una manera de hacer Europa”). Debo expresar mi agradecimiento a todos ellos y en especial a las investigadoras principales de los proyectos de Canarias, Lidia Pérez García de la Fundación Canaria de Investigación Sanitaria (FUNCANIS) y del proyecto del País Vasco, Amaia Bilbao del Instituto Kronikgune para la Investigación en Servicios de Salud. Los tres proyectos se consiguieron desde la Red de Investigación en Servicios Sanitarios y Enfermedades Crónicas (REDISSEC).

A Oxford University Innovation por proporcionarnos las versiones adaptadas al español de España de los cuestionarios Oxford Hip Score y Oxford Knee Score.

También quiero agradecer a Jesús, todo el tiempo que me ha dedicado en la corrección y mejora de este trabajo. Siempre quedarán en mi recuerdo todas esas mañanas de miércoles que pasamos revisando el trabajo, aprovechando mis salientes de guardia, algo que lo hace todavía más meritorio por tu parte.

No puedo dejar de mencionar a toda mi familia, a mis padres que me han permitido emprender este largo camino desde su origen, a mi pareja Gloria, como ejemplo de superación y capacidad de trabajo junto con su apoyo más personal; a mis hijos Elia y Antonio, por su comprensión, por el tiempo que no hemos pasado juntos, y su papel motivador para conseguir afrontar este difícil reto. Tampoco quiero olvidar a las personas que de algún modo estáis presentes (el tío Jesús, la Nona) y que sin duda seguro que os alegráis de mis logros personales y profesionales.

INDICE

AGRADECIMIENTOS.....	1
INDICE	5
LISTADO DE TABLAS Y FIGURAS	11
Listado de tablas.....	13
Listado de figuras	15
LISTADO DE ABREVIATURAS	17
1. RESUMEN.....	19
SUMMARY.....	25
2. INTRODUCCIÓN.....	29
2.1. La artrosis: definición, patogenia, factores de riesgo y manifestaciones clínicas	31
2.1.1 Definición de la artrosis.....	31
2.1.2 Patogenia de la artrosis.....	32
2.1.3 Factores de riesgo para la artrosis	35
2.1.4 Manifestaciones clínicas de la artrosis.....	45
2.2. La artrosis: epidemiología y carga de enfermedad.....	47
2.3. La artrosis de miembro inferior	51
2.4. Carga de enfermedad de la artrosis de miembro inferior.....	57
2.5. Medición de los resultados en salud desde la perspectiva del paciente.....	62
2.6. La calidad de vida relacionada con la salud	68
2.7. Herramientas de medida de la calidad de vida relacionada con la salud.....	72
2.8. Medición de la calidad de vida relacionada con la salud en pacientes con artrosis de miembro inferior	77
2.9. Validación de herramientas para medir la calidad de vida relacionada con la salud.....	87
3. JUSTIFICACIÓN	95
4. HIPÓTESIS Y OBJETIVOS.....	101
Hipótesis	103

Objetivos.....	105
5. MATERIAL Y MÉTODOS.....	107
5.1. Diseño del estudio.....	109
5.2. Ámbito.....	109
5.3. Población de estudio y criterios de selección.....	110
5.4. Tamaño muestral y selección de la muestra.....	111
5.5. Variables.....	112
5.5.1 Variables de resultado.....	112
5.5.2 Variables independientes.....	114
5.6. Fuentes de información.....	116
5.7. Análisis.....	117
5.7.1 Análisis descriptivo.....	117
5.7.2 Análisis inferencial.....	118
5.8. Aspectos éticos y legales.....	123
6. RESULTADOS.....	125
6.1. Estudio del impacto de la artrosis de miembro inferior en la CVRS.....	131
6.2. Validación del OHS.....	141
6.2.1 Aceptabilidad.....	142
6.2.2 Validez de contenido.....	142
6.2.3 Fiabilidad.....	144
6.2.4 Sensibilidad al cambio.....	149
6.2.5 Interpretabilidad.....	151
6.3. Validación del OKS.....	153
6.3.1 Aceptabilidad.....	154
6.3.2 Validez de contenido.....	154
6.3.3 Fiabilidad.....	156
6.3.4 Sensibilidad al cambio.....	161
6.3.5 Interpretabilidad.....	163
7. DISCUSIÓN.....	165
7.1 . Impacto de la artrosis de miembro inferior en la CVRS.....	170
7.1.1. Limitaciones.....	173
7.1.2. Implicaciones.....	174

7.2	Validación del Oxford Hip Score (OHS)	175
7.2.1.	Limitaciones.....	179
7.2.2.	Implicaciones.....	179
7.3.	Validación del Oxford Knee Score (OKS)	181
7.3.1.	Limitaciones.....	184
7.3.2.	Implicaciones.....	184
8.	CONCLUSIONES	187
9.	BIBLIOGRAFÍA	191
10.	ANEXOS	221
	Anexo 1. Versión adaptada al español (España) del Oxford Hip Score	223
	Anexo 2. Versión adaptada al español (España) del Oxford Knee Score.	225
	Anexo 3. Publicaciones	227

LISTADO DE TABLAS Y FIGURAS

Listado de tablas

Tabla 1. Características de la población incluida	128
Tabla 2. Percepción del estado de salud en los pacientes con artrosis de cadera y rodilla....	130
Tabla 3. Características de la muestra estudiada.....	132
Tabla 4. Distribución de las respuestas a los diferentes dominios del EQ-5D-5L en población general y con artrosis de miembro inferior (OAMI).....	134
Tabla 5. Características de los pacientes con OAMI incluidos en el estudio.....	136
Tabla 6. Modelos explicativos de las diferencias en las utilidades derivadas del EQ-5D-5L ...	138
Tabla 7. Modelos explicativos de las diferencias en las utilidades derivadas del EQ-5D-5L solo para pacientes con artrosis de miembro inferior.....	140
Tabla 8. Puntuaciones expresadas en los cuestionarios OHS, WOMAC y EQ-5D-5L.....	141
Tabla 9. Puntuaciones medias del OHS en los pacientes clasificados según los terciles en los cuestionarios WOMAC y EQ-5D-5L.	143
Tabla 10. Correlación (p de Pearson) entre las puntuaciones del OHS, las escalas del WOMAC y el EQ-5D-5L (índice de utilidad y escala EVA).....	144
Tabla 11. Análisis Factorial Exploratorio: extracción de componentes principales.....	145
Tabla 12. Análisis factorial exploratorio para los ítems del Oxford Hip Score (OHS).....	147
Tabla 13. Cambios en los cuestionarios OHS, EQ-5D-5L y WOMAC a los 6 meses de seguimiento en aquellos pacientes que referían cambios en su estado.	150
Tabla 14. Puntuaciones expresadas en los cuestionarios OKS, WOMAC y EQ-5D-5L.	153
Tabla 15. Puntuaciones medias del OKS en los pacientes clasificados según los terciles en los cuestionarios WOMAC y EQ-5D-5L.	155
Tabla 16. Correlación (p de Pearson) entre las puntuaciones del OKS, las escalas del WOMAC y el EQ-5D-5L (índice de utilidad y escala EVA).....	156

Tabla 17. Análisis Factorial Exploratorio (AFE), extracción de componentes principales autovalores para los ítems del Oxford Knee Score (OKS).-.....	157
Tabla 18. Análisis factorial exploratorio (AFE) para los ítems del Oxford Knee Score (OKS)..	159
Tabla 19. Cambios en los cuestionarios OKS, WOMAC y EQ-5D-5L a los 6 meses de seguimiento en aquellos pacientes que referían cambios en su estado.	162

Listado de figuras

Figura 1. Diagrama de flujo de la población incluida	127
Figura 2. Gráfico de sedimentación del Análisis Factorial Exploratorio (AFE) del cuestionario Oxford Hip Score (OHS).....	146
Figura 3. Análisis factorial confirmatorio para datos categóricos del cuestionario Oxford Hip Score (OHS). Se muestran los parámetros estandarizados y los errores estándar (flechas a la derecha).	148
Figura 4. Gráfico de sedimentación del Análisis Factorial Exploratorio (AFE) del cuestionario Oxford Knee Score (OKS).....	158
Figura 5. Análisis factorial confirmatorio para datos categóricos del cuestionario Oxford Knee Score (OKS). Se muestran los parámetros estandarizados y los errores estándar (flechas a la derecha).	160

Listado de abreviaturas

AIC: Akaike Information Bayesian

ACR: American College of Rheumatology

AFC: Análisis Factorial Confirmatorio

AFE: Análisis Factorial Exploratorio

AVAD: Año de Vida Ajustado por Discapacidad

AVD: Años de Vida con Discapacidad

AVP: Años de Vida Perdidos

BIC: Bayesian Information Criterion

CV: Calidad de Vida

CVRS: Calidad de Vida Relacionada con la Salud

ENSE: Encuesta Nacional de Salud Española

EQ-5D: European Quality of Life-5 Dimensions

EQ-5D-3L: European Quality of Life-5 Dimensions- 3 Level

EQ-5D-5L: European Quality of Life-5 Dimensions- 5 Level

EEM: Error Estándar de Medición

GWAS: Genome-Wide Association Study

HAGOS: Copenhagen Hip and Groin Outcome Score

HAQ: Health Assessment Questionnaire

HHR: Harris Hip Score

HOOS: Hip dysfunction and Osteoarthritis Outcome Score

HOS: Hip Outcome Score

HR-PRO: Health Related Patient Reported Outcomes

ICOAP: Intermittent and Constant Osteoarthritis Pain

iHOT : International Hip Outcome Tool

IKDC: International Knee Documentation Committee

IMC: Índice de Masa Corporal

KOOS: Knee Injury and Osteoarthritis Outcome Score

KSS: Knee Society Score

MCD: Mínimo Cambio Detectable

MDCI: Mínima Diferencia Clínicamente Importante

MMP: Metaloproteinasas

OA: Osteoartritis, artrosis

OAKHQOL: Osteoarthritis Knee and Hip Quality of Life

OAMI: Artrosis de Miembro Inferior

OARSI: Osteoarthritis Research Society International

OHS: Oxford Hip Score

OKS: Oxford Knee Score

OKS-PS: Oxford Knee Score Pain Subscale

OMS: Organización Mundial de la Salud

OARSI: Osteoarthritis Research Society International

SEM: Standard Error of Measure

SF-36: Short Form – 36

TCT: Teoría Clásica de los Test

TRI: Teoría de Respuesta al Ítem

VSG: Velocidad de Sedimentación Globular

WHOQOL: World Health Organization quality of life assessment

WOMAC: Western Ontario and McMaster Universities Osteoarthritis

1. RESUMEN

Introducción

La artrosis de miembro inferior (OAMI) es la primera causa de dolor y discapacidad funcional en la población mundial, suponiendo una pérdida en años de vida ajustados por discapacidad del 0,69% de todos los AVAD perdidos en la población mundial en 2010.

El impacto de la artrosis sobre la calidad de vida puede valorarse con herramientas que incorporan la perspectiva del propio paciente.

Por ello se pretende estudiar el impacto sobre la salud de la OAMI desde la perspectiva del paciente y evaluar el desempeño de nuevas herramientas de valoración de la CVRS, como el cuestionario Oxford Hip Score (OHS) o el Oxford Knee Score (OKS), específicas para estas condiciones.

Objetivos

Objetivo general:

Estimar el impacto de la artrosis de miembro inferior en la percepción de la CVRS en la población general y valorar las propiedades psicométricas de los instrumentos de medida específicos OHS y OKS, en sus versiones adaptadas al castellano.

Objetivos específicos

1. Evaluar el impacto de la artrosis de miembro inferior en la Calidad de Vida Relacionada con la Salud percibida por los sujetos que la sufren, comparándola con la situación de la población general.
2. Comparar las diferencias en la percepción de la CVRS entre sujetos que sufren artrosis de cadera o rodilla.
3. Estudiar potenciales características individuales y contextuales que se asocien con la percepción de la CVRS en pacientes con artrosis de miembro inferior.
4. Evaluar las características de fiabilidad y validez del cuestionario Oxford Hip Score, adaptado al castellano en nuestro medio.
5. Valorar la sensibilidad al cambio del cuestionario Oxford Hip Score, adaptado al castellano en nuestro medio.
6. Evaluar las características de fiabilidad y validez del cuestionario Oxford Knee Score, adaptado al castellano en nuestro medio.

7. Valorar la sensibilidad al cambio del cuestionario Oxford Knee Score, adaptado al castellano en nuestro medio.

Material y métodos

Se diseñó un estudio observacional longitudinal multicéntrico, con enfoque analítico, con un seguimiento de seis meses. Se incluyeron pacientes diagnosticados de artrosis de cadera y/o rodilla. Como grupo de comparación (objetivo 1) se escogió una muestra representativa de la población española ≥ 50 años, de la Encuesta Nacional de Salud Española 2011-12.

Como variable de resultado se utilizaron las puntuaciones derivadas del EQ-5D-5L (objetivos 1 a 3), la puntuación obtenida en la versión española del cuestionario específico OHS (objetivos 4 y 5) y la puntuación obtenida en la versión española del cuestionario específico OKS (objetivos 6 y 7).

Para dar respuesta a los objetivos 1 a 3 se construyeron Modelos Lineales Generalizados con variable dependiente el índice de utilidad atribuible al estado de salud descrito por el sujeto.

Para dar respuesta a los objetivos 4 a 7 se llevó a cabo un estudio de validación de un cuestionario evaluando la factibilidad, la validez, la fiabilidad, la sensibilidad al cambio y la interpretabilidad.

Resultados

Fueron incluidos en el estudio 758 pacientes, 361 con artrosis de cadera y 397 con artrosis de rodilla, procedentes de consultas de Atención Primaria y hospitalarias de Vizcaya, Madrid y Canarias.

La percepción del estado de salud y la CVRS expresada por los pacientes con OAMI fue significativamente peor que para la población general (0,532 frente a 0,924 en el índice de utilidades del EQ-5D-5L).

En los modelos explicativos se observa como la artrosis de cadera disminuye las utilidades en personas de similar edad, sexo, grupo social, estado de convivencia y carga de enfermedad crónica, en una media de 0,347 puntos. Para los pacientes que presentan artrosis de rodilla la disminución es de 0,295 puntos.

En ambos cuestionarios se descartaron “efecto suelo y efecto techo” y la capacidad discriminativa de ambos fue muy buena. Se observó una excelente fiabilidad test-retest en ambos, con coeficientes de correlación intraclase de 0,992 y 0,997 para el OHS y OKS respectivamente. La consistencia interna se evaluó mediante el alfa de Cronbach que se calculó en 0,928 para el OHS, y 0,920 para el OKS. El tamaño del efecto para la mejoría moderada en la CVRS percibida fue de 0,73 y 0,69, para OHS y OKS respectivamente. La mínima diferencia clínicamente importante (MDCI) y el mínimo cambio detectable (MCD) fue de 7,0 puntos y 5,5 para el OHS, y de 6,1 puntos y 4,38 para el OKS.

El análisis factorial exploratorio (AFE) muestra una estructura unidimensional en ambos cuestionarios.

Conclusiones

1. La artrosis de miembros inferiores (OAMI) se asocia a una enorme disminución en la utilidad atribuida al estado de salud, una vez excluida la influencia de otras variables que se relacionan con la Calidad de Vida Relacionada con la Salud (CVRS), superando con mucho los umbrales establecidos para la relevancia clínica.
2. La artrosis de rodilla empeora la percepción de CVRS en menor medida que la artrosis de cadera, una vez ajustada esta asociación por otras potenciales características que actúan como variables de confusión.
3. Cuando se comparan con la población general, las personas con artrosis de miembro inferior de mayor edad, las de grupos socioeconómicos más desfavorecidos, las que viven solas, las mujeres, las personas obesas y las que presentan más condiciones crónicas, suelen tener peor percepción del estado de salud. Las diferencias en percepción de CVRS entre personas ya diagnosticadas de artrosis de miembro inferior solo vienen explicadas por el tipo de artrosis, el sexo (peor percepción las mujeres), la presencia de otras enfermedades crónicas y, marginalmente, por la afectación bilateral.
4. El cuestionario Oxford Hip Score (OHS), en su versión en castellano, es una herramienta aceptable válida y fiable para medir la CVRS en pacientes con artrosis de cadera.

5. La capacidad de la versión en castellano del OHS para evaluar cambios con relevancia clínica a nivel individual, en sujetos con artrosis de cadera, es adecuada.
6. El cuestionario Oxford Knee Score (OKS), en su versión en castellano, es una herramienta aceptable, válida y fiable para medir la CVRS en pacientes con artrosis de rodilla.
7. La capacidad de la versión en castellano del OKS para evaluar cambios con relevancia clínica a nivel individual, en sujetos con artrosis de rodilla, es adecuada permitiendo distinguir con claridad los cambios potenciales del “ruido aleatorio” propio de cualquier instrumento de medida.

Introduction

Osteoarthritis of the lower limb (OALL) is the major cause of pain and functional disability in the world population and its impact in the form of disease burden is extraordinary, assuming a loss in disability-adjusted life years (DALYs) of 0.69% of all DALYs lost in the world population in 2010.

The impact of osteoarthritis on quality of life can be assessed with tools that incorporate the patient's own perspective.

The generic tools to assess Health-Related Quality of Life (HRQoL) have a high degree of development and use, but those instruments specific to a certain clinical condition, such as the Oxford Hip Score (OHS) questionnaire or the Oxford Knee Score (OKS), can be more useful to discriminate the situation of subjects with said disease and more sensitive to the changes that occur.

The use of these tools requires a process that ensures that they are useful for their purpose, so their psychometric properties must be evaluated.

These conditions lead us to find pertinent the study of the impact on health, from the patient's perspective, of lower limb osteoarthritis and the evaluation of new HRQL assessment tools, specific for this clinical condition

Objective

General objective:

Estimate the impact of lower limb osteoarthritis on the perception of Health-Related Quality of Life (HRQoL) in the general population and assess the psychometric properties of the specific measuring instruments OHS and OKS, in their versions adapted to Spanish.

Specific objectives

1. Evaluate the impact of lower limb osteoarthritis on the Health-Related Quality of Life perceived by the subjects who suffer it, comparing it with the situation of the general population, adjusting for sociodemographic characteristics and other clinical conditions.
2. Compare the differences in the perception of HRQoL among subjects suffering from hip or knee osteoarthritis.

3. Study potential individual and contextual characteristics that are associated with the perception of HRQL in patients with lower limb osteoarthritis.
4. Evaluate the reliability and validity characteristics of the Oxford Hip Score questionnaire, adapted to Spanish in our setting.
5. Assess the sensitivity to change of the Oxford Hip Score questionnaire, adapted to Spanish in our setting.
6. Evaluate the reliability and validity characteristics of the Oxford Knee Score questionnaire, adapted to Spanish in our setting.
7. Assess the sensitivity to change of the Oxford Knee Score questionnaire, adapted to Spanish in our setting.

Material and methods

A multicenter longitudinal observational study was designed, with an analytical approach, with a six-month follow-up. Patients diagnosed with hip and/or knee osteoarthritis were included. As a comparison group (objective 1), a representative sample of the 50-year-old Spanish population was chosen from the 2011-12 National Health Survey (NHS).

The main result variable used was the transformation into utilities of the generic questionnaire EQ-5D-5L (objectives 1 to 3), the score obtained in the adapted to Spanish for Spain version of the specific OHS questionnaire (objectives 4 and 5) and the score obtained in the adapted to Spanish for Spain version of the specific OKS questionnaire (objectives 6 and 7).

To respond to objectives 1 to 3, Generalized Linear Models were constructed, with the utility index attributable to the subject's perceived health status as the dependent variable.

To respond to objectives 4 to 7, a validation study of a questionnaire was carried out evaluating the different aspects of: feasibility, validity, reliability, sensitivity to change and interpretability.

Results

The study included 758 patients, 361 with hip osteoarthritis and 397 with knee osteoarthritis, coming from Primary Care and hospital consultations in Vizcaya, Madrid and the Canary Islands.

The perceived health status and HRQL expressed by patients with OALL was significantly worse than for the general population, with a loss of utility derived from the EQ-5D-5L of 0.392 points (0.532 vs 0.924).

In the explanatory models of the utilities, it is observed how hip osteoarthritis decreases the utilities in people of same age, sex, social group, state of coexistence and burden of chronic disease, by an average of 0.347 points. For patients with knee osteoarthritis, the decrease is 0.295 points.

In both questionnaires "Floor effect and Ceiling effect" were ruled out and the discriminative capacity of both was very good. The correlations with the WOMAC and the EQ-5D-5L utilities were strong with the scores of both questionnaires ($\rho > 0.7$). Excellent test-retest reliability was observed in both, with intraclass correlation coefficients of 0.992 and 0.997 for the OHS and OKS respectively. Internal consistency was evaluated using Cronbach's alpha, which was calculated at 0.928 for the OHS, and 0.920 for the OKS. The effect size for moderate improvement in perceived HRQL was 0.73 and 0.69, for OHS and OKS respectively.

The minimal clinically important difference (MCID) and the minimum detectable change (MDC) was 7.0 points and 5.5 for the OHS, and 6.1 points and 4.38 for the OKS.

The exploratory factor analysis (EFA) shows a one-dimensional structure in both questionnaires.

Conclusions

1. Osteoarthritis of the lower limbs (OALL) is associated with an enormous decrease in the utility attributed to the state of health, once the influence of other variables that are related to Health-Related Quality of Life (HRQOL) has been excluded. The impact produced by the OALL on HRQL measured in profits far exceeds the thresholds established for clinical relevance.
2. Knee osteoarthritis worsens the perception of HRQL to a lesser extent than hip osteoarthritis, once this association has been adjusted for other potential characteristics that act as confounding variables. These differences are close to the thresholds to which clinical relevance is attributed in the literature.

3. When compared to the general population, older people with lower limb osteoarthritis, those from more disadvantaged socioeconomic groups, those who live alone, women, obese people and those with more chronic conditions tend to have worse perception of health status. Differences in the perception of HRQoL between people already diagnosed with lower limb osteoarthritis are only explained by the type of osteoarthritis, gender (worse perception in women), the presence of other chronic diseases and, marginally, by bilateral involvement.
4. The Oxford Hip Score (OHS) questionnaire, in its version in Spanish, is an acceptable, valid and reliable tool to measure HRQL in patients with hip osteoarthritis.
5. The capacity of the Spanish version of the OHS to evaluate changes with clinical relevance at the individual level, in subjects with hip osteoarthritis, is adequate, clearly distinguishing potential changes from the “random noise” typical of any measurement instrument.
6. The Oxford Knee Score (OKS) questionnaire, in its Spanish version, is an acceptable, valid and reliable tool to measure HRQL in patients with knee osteoarthritis.
7. The capacity of the Spanish version of the OKS to evaluate changes with clinical relevance at the individual level, in subjects with knee osteoarthritis, is adequate, allowing the potential changes to be clearly distinguished from the “random noise” typical of any measurement instrument.

2. INTRODUCCIÓN

INTRODUCCIÓN

2.1. La artrosis: definición, patogenia, factores de riesgo y manifestaciones clínicas

2.1.1. *Definición de la artrosis*

La artrosis es una enfermedad heterogénea a la que se le asignan diversos nombres, como artropatía degenerativa y osteoartritis (OA). Es una artropatía caracterizada clínicamente por dolor y deterioro funcional. Aunque es frecuente la inflamación local de las articulaciones afectadas, no se asocia a un proceso inflamatorio general, a diferencia de otras artritis. Se considera la forma más frecuente de artritis.

La artrosis fue definida por la OMS en 1995 como un proceso degenerativo articular que se produce como consecuencia de trastornos mecánicos y biológicos que desestabilizan el equilibrio entre la síntesis y la degradación del cartílago articular, estimulando el crecimiento del hueso subcondral y que se asocia con la presencia de sinovitis crónica de intensidad leve (1).

Las definiciones formales de artrosis han evolucionado conforme ha avanzado el conocimiento de su fisiopatología.

La artrosis es en realidad una combinación compleja de cambios articulares, genéticos, metabólicos, biomecánicos y bioquímicos que afectan a la totalidad de la articulación y los tejidos que la rodean. La secuencia de la afectación de las estructuras articulares es variable, pero la característica común de la artrosis es un fracaso precoz de la remodelación del cartílago normal que ocasiona una degeneración condral en respuesta al estrés o una lesión (2).

El grupo de trabajo de la Osteoarthritis Research Society International (OARSI), ha definido la artrosis como “una afección de las articulaciones móviles, 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 proinflamatorias de la inmunidad innata” (3).

2.1.2. Patogenia de la artrosis

Sobre la patogénesis de la artrosis, aún no se conocen todos los aspectos sobre su aparición y desarrollo. Los investigadores han propuesto varios mecanismos. Uno de ellos es la degradación por proteasas de la matriz del cartílago (4,5). Las metaloproteinasas de la matriz (MMP) son enzimas que pueden descomponer el colágeno, que se encuentra en los espacios entre las células de los tejidos; de forma que promueven la degradación de la matriz del cartílago e inducen la apoptosis de los condrocitos articulares (6). Las metaloproteinasas se han clasificado en gelatinasas (MMP-2, MMP-9), colagenasas (MMP-1, MMP-8, MMP-13) y estromelisinias (MMP-3, MMP-7, MMP-10, MMP-11) (7). De estas metaloproteinasas, la MMP-3 es la más conocida por su papel en la degradación de proteoglicanos y también en la activación de procólagenasa.

La expresión y actividad de las metaloproteinasas de matriz local está regulada por metaloproteinasas de matriz derivadas de tejido in vivo. El desequilibrio dinámico entre los dos puede estar en el origen de la aparición de artrosis (8).

El otro mecanismo fisiopatogénico que se describe en la artrosis se enmarca en la teoría de las citocinas. Las citocinas se consideran factores importantes en la aparición y el desarrollo de la artrosis, como la interleucina 17, (9) y el factor de necrosis tumoral (TNF- α) (10) que aumentan significativamente en la artrosis, lo que conduce al daño del tejido del cartílago articular y causa inflamación y dolor. Se ha demostrado que los factores de crecimiento, como el factor de crecimiento transformante beta (TGF- β), protegen contra la degradación del cartílago articular (11). El desarrollo de la artrosis se explicaría como un desequilibrio entre factores inflamatorios y protectores del cartílago.

Como decimos desconocemos en parte los mecanismos etiopatogénicos de la artrosis, si bien desde un punto de vista clásico, ésta se puede clasificar en primaria o idiopática, en ausencia de un antecedente de lesión o de otra artropatía; y secundaria, si se debe a un proceso concreto que ocasiona o empeora el desarrollo de artrosis.

La artrosis primaria, o idiopática, puede clasificarse a su vez en localizada o generalizada. La artrosis localizada suele afectar a una sola articulación; la artrosis generalizada consiste en la afectación de tres o más articulaciones.

Las causas de artrosis secundaria son numerosas e incluyen trastornos metabólicos (como la artritis por cristales, acromegalia, ocronosis, diabetes, hemacromatosis, hiperparatiroidismo, enfermedad de Wilson, síndrome de Ehlers-Danlos o la enfermedad de Gaucher), trastornos mecánicos (como las displasias epifisarias, enfermedad de Perthes, displasia del desarrollo de la cadera, disimetría de extremidades inferiores, síndromes de hipermovilidad o la osteonecrosis), lesiones traumáticas previas (como traumatismos articulares, fracturas intraarticulares, cirugía articular previa o la artropatía de Charcot), y enfermedades articulares inflamatorias (artritis reumatoide, artropatía cristalina o antecedentes de artritis séptica).

Al aumentar el número de factores de riesgo que se conocen para desarrollar artrosis a nivel local, y al ampliar la gama de factores asociados tanto genéticos, como biomecánicos o medioambientales, la división entre artrosis primaria y secundaria es más difusa.

Por lo que puede ser más útil pensar en la artrosis como una enfermedad crónica, a través de la cual la genética, los antecedentes de una lesión y de otros daños articulares, los factores mecánicos y el entorno psicosocial actúan en la articulación provocando un deterioro o desgaste de la superficie articular.

Es probable que los mecanismos que conducen al deterioro de la articulación sean multifactoriales, como consecuencia de numerosos factores individuales. Estos incluyen daño oxidativo, adelgazamiento del cartílago, debilitamiento de los músculos y reducción de la propiocepción. Además, los mecanismos celulares básicos que mantienen la homeostasis tisular disminuyen con el envejecimiento, lo que conduce a una respuesta inadecuada al estrés o lesión articular y la destrucción y pérdida de tejido articular resultante (12) .

Una clasificación más actual de la artrosis, se basa en la información epigenética de los pacientes, la cual está modulada por diferentes factores ambientales, y nos serviría para definir una clasificación de diferentes fenotipos de la artrosis.

Un fenotipo puede entenderse como las características observables de un individuo que resultan de factores genéticos, combinados con factores ambientales. Se considera que los subgrupos de pacientes que tienen características clínicamente similares representan un fenotipo.

INTRODUCCIÓN

La finalidad de la clasificación en fenotipos es identificar mejor a los individuos con mayor riesgo de progresión de la artrosis (fenotipado pronóstico) o de estratificar en subpoblaciones de pacientes con más probabilidad de responder a un tratamiento dirigido, o a medidas de prevención estratégicas (fenotipado prescriptivo).

En estudios de revisión sistemática recientes de los fenotipos de la artrosis se han descrito seis tipos que incluían: 1) fenotipo de dolor crónico con sensibilización central; 2) fenotipo inflamatorio; 3) fenotipo del síndrome metabólico; 4) fenotipo del metabolismo óseo y cartilaginoso; 5) fenotipo mecánico (mala alineación); y 6) fenotipo de enfermedad articular mínima (13).

No está claro si los seis fenotipos propuestos son entidades distintas o si hay superposición entre ellos. Tampoco es posible afirmar que la mayoría de la población con artrosis encajaría en al menos una de las categorías fenotípicas propuestas.

Otro concepto que ha surgido de los estudios de fenotipado en enfermedades crónicas es el término endotipo. A diferencia de los fenotipos, que se basan en características clínicas que no están necesariamente conectadas a un mecanismo fisiopatológico establecido de la enfermedad, un endotipo es un subtipo de enfermedad definida funcional y patológicamente por un mecanismo molecular (14).

Los diferentes fenotipos clínicos de artrosis pueden consistir en endotipos moleculares superpuestos que pueden identificarse por la presencia de células específicas o moléculas de biomarcadores, medidos en suero, plasma, orina o líquido sinovial, que es el líquido tisular proximal más adecuado para evaluar la taxonomía molecular y celular de los subtipos de artrosis (13).

Hasta la fecha, no se ha encontrado que un solo marcador sea suficiente para el diagnóstico o pronóstico en la artrosis.

Una mejor comprensión de los endotipos moleculares subyacentes nos permitirá definir los fenotipos clínicos con mayor claridad y desarrollar paneles de biomarcadores que puedan predecir la progresión de la enfermedad y determinar qué pacientes pueden tener una mejor capacidad de reparación articular. Incluso podemos distinguir entre los pacientes que se beneficiarán de una cirugía de reemplazo total de articulaciones anterior y aquellos que pueden conservar su articulación durante más tiempo.

2.1.3. Factores de riesgo para la artrosis

Volviendo a una concepción más clásica de la enfermedad, los factores de riesgo de aparición de artrosis, se pueden clasificar en factores personales y factores de riesgo articular.

Dentro de los factores de riesgo personales, podemos identificar factores sociodemográficos (edad, sexo y etnia), asociados al estilo de vida (tabaquismo y alimentación), genéticos y biomarcadores, y los factores de riesgo clínicos (densidad mineral ósea y comorbilidades). En cuanto a los factores de riesgo articular, constituyen un grupo de factores más conocidos y estudiados que se relacionan con el concepto de fenotipo mecánico de la enfermedad.

La edad es el principal factor de riesgo en la artrosis en sus diferentes localizaciones. En efecto, la prevalencia y la incidencia de la artrosis radiológica aumentan de forma considerable con la edad (15), aunque la tendencia no es lineal. Entre los 50-75 años se observa un fuerte aumento de la incidencia, que es más limitada después de los 75 años (16).

El sexo femenino se ha asociado a una mayor prevalencia y una mayor gravedad de la artrosis, en particular de la rodilla (17). encontrándose riesgos de padecerla de hasta casi un 70 % superiores en mujeres (16), mientras que no se han observado diferencias significativas por sexo para la coxartrosis (12).

La prevalencia de artrosis varía entre los grupos étnicos. La coxartrosis y la artrosis de los dedos son menos frecuentes en los chinos que en las personas de raza blanca. De forma paradójica, las mujeres chinas tienen una mayor prevalencia de gonartrosis, lo que puede explicarse por una carga excesiva de la articulación de la rodilla debido a la costumbre de estar en cuclillas. (18). Otro estudio ha demostrado que la prevalencia de la coxartrosis en las mujeres estadounidenses de raza negra y blanca era similar, pero esta prevalencia era ligeramente superior en los varones de raza negra (21%) respecto a los de raza blanca (17%) (19).

De entre los factores asociados al estilo de vida destaca la obesidad en sus diferentes grados. Los individuos con sobrepeso, definido por un índice de masa corporal (IMC) de 25-30 kg/m² tienen un mayor riesgo para desarrollar artrosis, siendo esta asociación más evidente para la rodilla (16). Este exceso de riesgo sigue aumentando a medida que aumenta el IMC, siendo más pronunciado para la artrosis de rodilla en sujetos con obesidad grado 2 (> 30 kg/m²); estos sujetos tienen 4,7

INTRODUCCIÓN

veces más probabilidades de desarrollar artrosis de rodilla en comparación con los sujetos con peso normal (20). Otros estudios también han demostrado una relación dosis-respuesta entre la obesidad y el riesgo de gonartrosis. Por cada aumento de 5 unidades del IMC, se asocia un aumento de riesgo de gonartrosis del 35%, y esta asociación es mayor en las mujeres. La obesidad también se asoció significativamente con la artrosis de la mano (OR 2,59; 1,08–6,19), pero no con la artrosis de cadera (OR 1,11; 0,41–2,97) (21).

El consumo de tabaco como factor de riesgo para el desarrollo de artrosis muestra resultados contradictorios en los estudios publicados. En un metaanálisis reciente sobre el tabaquismo y la artrosis en manos, rodillas, caderas y columna vertebral donde se resumieron los datos de 48 estudios publicados, se establece que existe una relación inversa entre el tabaquismo y la artrosis (OR= 0,87; IC del 95%, 0,80 a 0,94). Este efecto protector era mayor cuando no se ajustaba el IMC (22). Pero no todos los autores consideran probado este efecto protector pues puede haber sesgos de selección en los estudios realizados en entornos hospitalarios donde los sujetos de control tienen afecciones relacionadas con el tabaquismo (23).

La influencia de la dieta en el desarrollo de la artrosis es un campo de investigación de gran interés. Los niveles plasmáticos de las vitaminas D, K, E y C han sido objeto de múltiples estudios. Algunos estudios demuestran una asociación entre la deficiencia de vitamina D y la progresión de los síntomas de gonartrosis (24,25). Los participantes con niveles bajos de vitamina D (25-hidroxitiamina D) tenían incrementado el riesgo de progresión de la artrosis de rodilla en más del doble (24). El uso constante de suplementos de vitamina D durante 4 años se asoció con una progresión significativamente menor del desgaste del cartílago, medido mediante resonancia magnética (26).

La vitamina K también se ha contemplado como elemento protector frente a la artrosis, en su papel de activadora de las proteínas que contienen γ -carboxiglutamato que regulan negativamente la calcificación. La calcificación del cartílago se ha asociado con una disminución de la resistencia a la compresión y un aumento de la fricción del cartílago. Por lo tanto, la calcificación del cartílago junto con una mayor expresión del factor de necrosis tumoral alfa e interleucinas de macrófagos inducida por partículas de desgaste y rotura podría conducir a una mayor degradación del cartílago articular (27).

La mayoría de los estudios actuales atribuyen los efectos condroprotectores de la vitamina E a sus efectos antioxidantes. Los efectos de la suplementación con vitamina E en el retraso de la

progresión de la artrosis en los pacientes aún son discutibles debido a los resultados heterogéneos encontrados (28).

La vitamina C es un nutriente esencial para la biosíntesis de los componentes de la matriz de colágeno que forma el cartílago articular (29). En algunos estudios de cartílago animal, se ha demostrado que la vitamina C es condroprotectora (29), aunque niveles elevados de vitamina C circundante, no se asocian a protección contra la artrosis radiográfica de rodilla incipiente, e incluso pueden estar asociados con un mayor riesgo de gonartrosis (26,30).

También se ha estudiado el potencial papel de otros nutrientes, como los ácidos grasos en la artrosis. Los ácidos grasos poliinsaturados omega-3 disminuyen la producción de eicosanoides inflamatorios, citoquinas y especies reactivas de oxígeno, así como la expresión de las moléculas de adhesión. Además producen mediadores antiinflamatorios denominados resolvinas. (31,32). Los estudios de intervención en humanos con suplementos de ácidos grasos poliinsaturados omega-3 parecen asociarlos a un efecto beneficioso sobre el dolor y la función y pueden estar asociados con un menor daño estructural. Por el contrario, se ha observado un efecto adverso de los ácidos grasos saturados sobre la artrosis (33).

Numerosos estudios se han interesado también por el papel de los complementos alimentarios sobre la aparición y la progresión de la artrosis. Entre ellos destaca la cúrcuma que se ha formulado en diversas presentaciones (34). La cúrcuma es una hierba medicinal con una larga historia de uso como tratamiento para afecciones inflamatorias en China y el sudeste asiático (35). Algunos estudios aportan alguna evidencia positiva sobre la reducción de los síntomas de la artritis, fundamentalmente sobre el dolor y la inflamación, con el tratamiento con extractos de cúrcuma estandarizados (generalmente 1000 mg / día de curcumina) entre 8 y 12 semanas (36).

Se ha postulado la existencia de un riesgo genético para la artrosis. En trabajos sobre gemelos, que permiten obtener una estimación de la heredabilidad (37), se ha estimado esta carga genética en un 40% para la artrosis de rodilla, 60% para la artrosis de cadera, 65% para la mano y aproximadamente 70% para la columna (38).

Se ha demostrado que las modificaciones de la expresión génica de factores de transcripción específicos, de citocinas pro o antiinflamatorias, de proteinasas de matriz y de proteínas de matriz extracelular en el cartílago articular, pueden estar implicadas en el desarrollo de la artrosis.

INTRODUCCIÓN

Diversos estudios han puesto de relieve la importancia de la regulación epigenética de la expresión de los genes en el desarrollo de la artrosis (39).

La investigación sobre la genética molecular de la artrosis se ha visto reforzada sustancialmente en los últimos años mediante la implementación de estudios de asociación de genoma completo (de las siglas en inglés GWAS, Genome-wide Association Study), exploraciones de todo el genoma que han revelado una gran cantidad de nuevos loci de riesgo asociados con la enfermedad (Rice, Beier, Young, & Loughlin, 2020).

EL GDF-5 fue el primer locus identificado a nivel de estudios de asociación de genoma completo (GWAS) en europeos y asiáticos. El GDF-5 es un miembro de la familia de las proteínas morfogenéticas del hueso (BMP) que tiene una participación clave en el desarrollo del esqueleto, la condrogénesis, la formación de las articulaciones sinoviales, y también en los procesos de reparación del hueso y del cartílago. La asociación entre el GDF-5 y la artrosis se observó para la cadera y en menor medida con la artrosis de rodilla (40). El segundo locus que se encontró asociado a nivel de estudios de genoma completo (GWAS) en europeos es el locus del cromosoma 7q22. Es un locus muy extenso que ocupa unas 500 kb, y 6 genes diferentes: PRKAR2B, HBP1, COG5, GPR22, DUS4L y BCAP29; que se asoció con la artrosis de rodilla (41). Además se han identificado varias regiones en los cromosomas 2, 3, 4, 6, 7, 11, 16 y cromosoma X, en las que supuestamente se localizan genes de susceptibilidad para la artrosis. La asociación más coherente se ha señalado entre el cromosoma 2q13-32 y fenotipos específicos de la artrosis. En los cromosomas 6 y 16 también se han identificado locus implicados en la artrosis de los dedos y de la cadera (38).

Un amplio estudio reciente de asociación pangenómica ha suministrado pruebas evidentes de que un nuevo locus en el cromosoma 5p13, cerca de LSP1P3 se asociaba con el desarrollo de artrosis de rodilla radiológica. En este estudio también se ha confirmado la participación de los genes GDF y FTO en el desarrollo de gonartrosis radiológica (42).

De forma sintética y agrupando los resultados de los diversos estudios de genoma completo (GWAS), se observa que algunos de los genes implicados en el proceso de artrosis de rodilla son el DVWA, GLN3/GLT8D1, BTNL2, COG5/GPR22/DUS4L/HBP1, MCFL2, HLA-DQB1, TP63, FTO y SUPT3H / RUNX2. Y los genes implicados en la artrosis de cadera el DOT1L, GLN3/GLT8D1, ASTN2, FILIP1/SEN6, PTHLH, CHST11, SUPT3H/RUNX2, NCOA3, IGF3BP3 (43). No obstante, el número de loci es relativamente bajo en comparación con otras enfermedades complejas. Se sospecha que

uno de los motivos que explican este fenómeno es la heterogeneidad de fenotipos de la artrosis, que provocan que esta gran variabilidad reduzca el poder estadístico de los estudios de asociación. Otra forma de aumentar el poder de detección de nuevas variantes, es centrar los estudios genéticos en el estudio de endofenotipos, que son fenotipos más próximos a la etiología biológica de la enfermedad, que la enfermedad con su espectro de signos y síntomas. Se considera que los endofenotipos permiten obtener grupos de pacientes más homogéneos y tienen una relación más directa con los factores genéticos (37).

En el abordaje de los factores de riesgo para desarrollar artrosis, también se ha demostrado el papel de los marcadores bioquímicos como marcadores de la artrosis (44).

Entre los marcadores bioquímicos de destrucción de la matriz, el telopéptido C-terminal urinario del colágeno de tipo II (uCTX-II) ha sido el más estudiado, mostrando un incremento de la incidencia radiográfica de gonartrosis y un incremento de la incidencia de artrosis de cadera en más de cuatro veces (45). También se ha asociado en múltiples estudios con la progresión de la artrosis de rodilla y cadera (46).

Las metaloproteasas de matriz sérica (MMP) implicadas en la degradación del colágeno tipo II (C2M) presentan una asociación significativa en la incidencia y la progresión de la artrosis de rodilla (45). Otros marcadores de destrucción de la matriz como el telopéptido C-terminal urinario del colágeno tipo I (uCTX-I) y el telopéptido alfa C del colágeno tipo I (α -CTX) también se relacionaron con la progresión de la artrosis de cadera y rodilla, si bien cuenta con pocos estudios que respalden esta relación (47,48).

Entre los biomarcadores para la producción y diferenciación de matrices, solo la proteína oligomérica de la matriz del cartílago (COMP) cuenta con un número suficiente de estudios hasta el momento. Los niveles séricos elevados de esta proteína de la matriz extracelular no colágena se asocian con mayor riesgo de incidencia de artrosis de rodilla y/o cadera (45,49–51). Entre los muchos biomarcadores de inflamación, la proteína C reactiva (PCR) y el ácido hialurónico sérico son los más estudiados hasta el momento. El papel de los niveles de proteína C reactiva, como predictor de la incidencia y progresión de la artrosis de rodilla y cadera es incierto (46). El ácido hialurónico sérico se correlaciona positivamente con la progresión de la artrosis de rodilla (52,53) o de cadera (47), pero la valoración de estos biomarcadores debe ajustarse por la edad, el sexo y la etnia de los pacientes (54).

INTRODUCCIÓN

No obstante varias revisiones han proporcionado una descripción general de las publicaciones sobre el rendimiento de los marcadores bioquímicos en la artrosis primaria de rodilla y cadera (55–58) aplicando los criterios de consenso BIPED (59), y llegaron a la conclusión de que ninguno de los marcadores bioquímicos discriminaba lo suficiente como para permitir el diagnóstico y el pronóstico de la artrosis.

Entre los factores de riesgo clínico para desarrollar artrosis, señalamos la densidad mineral ósea y las comorbilidades. Hay una relación entre la densidad mineral ósea y la artrosis. Los estudios epidemiológicos han demostrado que las personas con una masa mineral ósea muy elevada, presentan una mayor prevalencia de artrosis radiológica de grandes articulaciones, en comparación con las poblaciones de control (60,61), e incluso se ha observado un aumento de la prevalencia en la cirugía de reemplazo articular (61)

El ajuste por el IMC atenuó la asociación entre la densidad ósea elevada y la artrosis de rodilla en aproximadamente un 50%, lo que sugiere que la asociación en la rodilla está mediada en parte por el IMC, situación que no ocurre en el caso de la artrosis de cadera (60).

El papel de las comorbilidades en el desarrollo y progresión de la artrosis está íntimamente relacionado con el concepto más actual de fenotipos de artrosis, concretamente con el fenotipo del síndrome metabólico.

La artrosis asociada al síndrome metabólico, resultante de la asociación de enfermedades metabólicas (obesidad, diabetes y resistencia a la insulina, dislipidemia e hipertensión) y la artrosis, ha sido recientemente estudiada, y nos conduce al concepto de regulación sistémica de la articulación (62). Epidemiológicamente ha sido demostrado que existe un relación entre síndrome metabólico y la artrosis. En un estudio de cohorte estadounidense con más de 7.000 personas, se ha demostrado que la prevalencia del síndrome metabólico era dos veces más elevada en los pacientes artrósicos que en la población general sin artrosis (62).

La obesidad afecta a los tejidos de las articulaciones no solo por el aumento de la carga que induce una tensión mecánica crónica en las articulaciones de las extremidades inferiores, sino también porque altera la homeostasis conjunta debido a un aumento sistémico de la inflamación que implican adipocinas, citocinas y ácidos grasos libres (63). Las adipocinas más estudiadas son la leptina, visfatina y resistina, y se asocian con efectos proinflamatorios y prodegradantes en los

tejidos articulares, mientras que la adiponectina, con una tasa sérica disminuida en pacientes obesos, parece tener un papel protector en la artrosis (64).

La diabetes tipo 2 o una glucemia elevada se asocian al desarrollo y a la progresión de la artrosis. El mecanismo que explica esta asociación viene determinado porque exponen a los tejidos articulares a niveles crónicamente altos de glucosa, que inducen la liberación de alto estrés oxidativo, citocinas y enzimas proteolíticas; así como la acumulación de productos finales de glicación avanzada (AGE) en los tejidos articulares, que estarían implicados en el proceso de degradación de la matriz cartilaginosa y en la activación de los condrocitos (65).

La hipercolesterolemia y la hipertrigliceridemia se han asociado con un mayor riesgo de artrosis, mientras que los niveles elevados de lipoproteínas de alta densidad (HDL) parecen proteger contra las lesiones tempranas de la médula ósea de la rodilla como se observa en estudios con resonancia magnética (66,67). La dislipidemia se asocia con niveles elevados de lipoproteínas de baja densidad de alta oxidación (ox-LDL), que aumenta la formación ectópica ósea mediante la activación del factor de crecimiento transformante- β (TGF- β) y promueve la inflamación de la sinovial debido a la activación de los macrófagos facilitando el desarrollo de artrosis (63).

Por último, la hipertensión se asocia con un aumento significativo de la incidencia de artrosis. Además la presencia de una gonartrosis radiológica aumentaba el riesgo de hipertensión y de dislipidemia (68).

Una vez analizados los factores de riesgo individuales para el desarrollo de artrosis, nos encontramos con un segundo grupo de factores, los factores de riesgo articular. Constituyen un grupo de factores más conocidos y estudiados que se imbrican con el concepto de fenotipo mecánico de la enfermedad.

Los antecedentes de traumatismos o lesiones osteoligamentosas, determinadas actividades físicas, riesgos ocupacionales y las alteraciones en la morfología y/o alineación articular se han asociado con la incidencia y progresión de la artrosis.

Las personas con una lesión previa de rodilla tienen un riesgo más elevado a desarrollar artrosis sintomática de rodilla, si bien el grado de heterogeneidad entre los hallazgos es considerable (16).

INTRODUCCIÓN

En la cadera, las lesiones traumáticas también aumentan el riesgo a desarrollar artrosis, pero la incidencia global de las mismas es muy inferior a las lesiones localizadas en la rodilla (69).

La lesión del ligamento cruzado anterior (LCA) es un factor de riesgo para la aparición temprana y la progresión acelerada de la artrosis, y la reconstrucción del ligamento cruzado anterior reduce, pero no elimina, el riesgo de desarrollar cambios degenerativos (70). La técnica elegida para la reconstrucción del ligamento cruzado anterior parece tener influencia en la incidencia para el desarrollo de artrosis. Las técnicas de reconstrucción con un solo túnel transtibial, no logran colocar el túnel femoral dentro de la huella del ligamento cruzado anterior nativo (71,72), lo que provoca una restauración más imperfecta de la cinemática articular en comparación con la reconstrucción anatómica, a través de un túnel femoral independiente (73,74). En un estudio reciente se estimó que la artrosis radiográfica se diagnosticó en aproximadamente el 23-36% de las rodillas operadas después de reconstrucciones anatómicas, y el 44-61% de las rodillas después de reconstrucciones no anatómicas, en un seguimiento medio de aproximadamente 15 años (75).

Una lesión de menisco o una lesión combinada que afecta tanto al ligamento anterior cruzado como al menisco se asocia con probabilidades seis veces mayores para el desarrollo de artrosis en comparación con una rodilla no lesionada (76). Por otra parte, el riesgo de desarrollar artrosis femorotibial es más elevado en las pacientes intervenidos mediante meniscectomía total frente a la meniscectomía parcial, siendo la cantidad de menisco resecado el principal factor predictivo de artrosis a largo plazo (77), aunque la meniscectomía parcial también se asocia a un mayor riesgo de empeoramiento de las lesiones cartilaginosas (78).

Los estudios sobre la influencia de las tareas profesionales en la aparición de artrosis son numerosos. Los trabajos físicamente exigentes se han asociado fuertemente con mayor riesgo de artrosis de la cadera y de rodilla, aunque los estudios observacionales que se han hecho están sometidos a diferentes tipos de sesgo (79).

Uno de los factores de riesgo ocupacionales más comunes para la artrosis de rodilla es la gran carga de trabajo físico (80). Otros factores de riesgo incluyen la exposición frecuente a varios factores estresantes biomecánicos como doblar la rodilla, arrodillarse o ponerse en cuclillas, estar de pie durante muchas horas (más de 2 horas al día), subir escaleras regularmente, levantar objetos pesados (por encima de 10 kg), saltos y vibraciones (38).

Encontramos múltiples estudios que nos presentan listados de trabajos (albañiles, carpinteros, fontaneros, agricultura, pesca, minería, etcétera) que requieren de un esfuerzo físico importante, y en las condiciones anteriormente descritas, que se asocian con el desarrollo de artrosis de rodilla de forma significativa (38).

En lo que respecta a la artrosis de cadera y su posible asociación con determinadas profesiones; existen pocos estudios, en comparación con la rodilla. La carga de trabajo físico pesado o el trabajo manual que implica cargas pesadas, al igual que ocurre con la rodilla, se identificó como un factor de riesgo ocupacional para la artrosis de cadera (69). El trabajo recurrente con el manejo de cargas superiores a 20 kilogramos se asoció con un riesgo 1,8 veces superior para el desarrollo de artrosis de cadera en ambos sexos; excepto en los jóvenes menores de 30 años. El tiempo necesario para desarrollar artrosis de cadera, estableciendo el diagnóstico mediante criterios clínicos y radiológicos fue superior a los 12 años en todos los grupos de edad (81). Respecto a las posiciones mantenidas que se asocian con mayor estrés biomecánico para la cadera, se encuentra el gateo o el permanecer de pie realizando un trabajo pesado (82). La sedestación durante más de dos horas al día se presenta como un factor protector (83).

Sin embargo, realizar una actividad física de forma habitual y sin cargas de peso, podría ser beneficioso al reforzar los músculos periarticulares y contribuir de esta manera a estabilizar la articulación. Los estudios en población general han demostrado que los niveles habituales de actividad física no se asocian a una artrosis sintomática o radiológica incidente. En cambio, los niveles más intensos de actividad física o los deportes de competición aumentarían el riesgo de artrosis. Pero los factores de riesgo parecen ser otros, además de la sobrecarga mecánica para la articulación (84,85), pues las lesiones previas constituirían el factor principal de riesgo de desarrollar artrosis de miembro inferior asociado a algunos deportes (86).

La anatomía o la morfología articular podrían contribuir al riesgo para desarrollar artrosis, debido a que la distribución de las cargas biomecánicas en la articulación depende de su morfología.

Otro factor a tener en cuenta son las alteraciones en los ejes de carga, dado que la desalineación en la extremidad inferior puede contribuir a generar una sobrecarga de la articulación, especialmente de la rodilla. La alineación de la rodilla es un claro predictor de la progresión de la artrosis de rodilla (87), sin embargo su influencia en la incidencia para el desarrollo de artrosis sigue siendo incierta (88). La mayor parte de la mala alineación en el plano frontal en pacientes con artrosis de rodilla es en varo. Esta alteración en varo, se mantiene durante la fase de apoyo

INTRODUCCIÓN

bipodal y durante la marcha, y puede valorarse fácilmente en estudios radiológicos de las extremidades inferiores en carga. Se ha demostrado que la mala alineación tibiofemoral en varo es un factor de riesgo extremadamente relevante para la progresión de la enfermedad, pero no ocurre lo mismo en las deformidades en valgo donde esta relación no es tan clara, y muestra resultados contradictorios para los investigadores (87,89).

En posición estática la deformidad en valgo puede mantenerse en la extremidad, pero durante la fase de apoyo, el vector de fuerza que se extiende desde el centro de presión del pie hasta la pelvis, pasa medial a la rodilla en muchos pacientes con valgo, y esto genera un momento de fuerza en varo. Así las rodillas con deformidades en valgo leves o moderadas no estarían en riesgo de desarrollar una artrosis tibiofemoral lateral o favorecer su progresión (87). Estos resultados son contradictorios a los presentados por otros investigadores, donde en las rodillas con artrosis establecida, incluso grados leves de alineación en valgo (1,1 a 3º en valgo) se asocian con un aumento sustancial de las probabilidades de progresión radiográfica de la artrosis femorotibial lateral, siendo la probabilidad mayor conforme aumenta la deformidad en valgo. Y en las rodillas sin alteraciones radiológicas de artrosis, la alineación en valgo por encima de 3º se asocia con daño del cartílago y por tanto con un riesgo mayor para desarrollar artrosis femorotibial lateral incipiente (89,90).

Las alteraciones en la morfología anatómica de la articulación de la cadera y su contribución al riesgo para desarrollar artrosis, han sido muy estudiadas en dicha articulación. Parámetros como la relación cabeza-cuello femoral, la esfericidad de la cabeza femoral, la cobertura acetabular, la anteversión acetabular y del cuello femoral pueden afectar a la patogenia de la artrosis (91). Las personas con defectos de cobertura de la cabeza femoral (displasia de cadera) y las que presentan un choque femoro-acetabular tipo "cam" o en "leva", causado por un atrapamiento de una cabeza femoral anormal, por la existencia de una prominencia ósea en la transición cabeza-cuello, duplican el riesgo de desarrollo coxartrosis. Mientras que el pinzamiento femoroacetabular tipo "pincer" o "en tenaza", causado por una anomalía acetabular, generalmente por sobrecobertura por retroversión acetabular no aumenta el riesgo de artrosis de cadera (92).

2.1.4. *Manifestaciones clínicas de la artrosis*

Las manifestaciones clínicas de la artrosis son muy variadas. 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.

La artrosis puede manifestarse durante un período prolongado con anomalías en el tejido musculoesquelético a un nivel molecular, pero clínicamente cursar de forma asintomática, precediendo las alteraciones anatómicas a las manifestaciones clínicas durante años o incluso décadas. Por tanto, la artrosis muy temprana se caracterizaría por un estado patológico asintomático (2).

La artrosis clínica se manifiesta con dolor en las articulaciones, que empeora con la actividad, con rigidez matutina breve (menor de 30 minutos), y dolor y rigidez con el reposo. La rigidez matutina es menos intensa por lo general y más breve que la observada en personas con artropatías inflamatorias sistémicas.

Es común, que las características radiográficas que definen a la artrosis se encuentren a menudo en articulaciones de personas sin síntomas. En este sentido, la presencia de las alteraciones radiológicas características de la artrosis pueden considerarse factores de riesgo para el desarrollo de la artrosis clínica.

Atendiendo a la disociación entre la clínica y las alteraciones anatómicas presentes en los estudios radiológicos, se distinguen, para su caracterización epidemiológica, cuatro definiciones diferentes de la artrosis: artrosis referida por el paciente, artrosis radiológica, artrosis sintomática (presencia de síntomas y de signos radiológicos) y artrosis clínica (presencia de síntomas, de signos clínicos y de signos radiológicos). La definición que combina las características clínicas y radiológicas es la más extendida en salud pública, ya que refleja de forma más precisa la carga real de la enfermedad.

La presencia y gravedad de la artrosis se ha definido en la investigación clínica por patrones radiológicos. En este sentido la escala de Kellgren y Lawrence (93) es un índice compuesto en el que se tienen en cuenta los osteofitos y el pinzamiento de la interlínea articular. Esta escala se puede usar para la artrosis de la mano y de la cadera, aunque respecto a la rodilla se ha

INTRODUCCIÓN

desarrollado para definir la artrosis femorotibial (y no para la artrosis de la articulación femorrotuliana).

Además, los criterios del American College of Rheumatology también se pueden usar para definir, de manera estandarizada, los casos de artrosis clínica de la rodilla (94), cadera (95) y de los dedos (96).

2.2. La artrosis: epidemiología y carga de enfermedad

La carga de enfermedad representa el impacto de un problema de salud medido por su coste financiero, la mortalidad, la morbilidad u otros indicadores. Para captar el impacto de la enfermedad en una sociedad, tanto en forma de muerte prematura como de discapacidad, se necesita una medida estandarizada que permita medir cómo afectan a esa sociedad diversas enfermedades y compararlas entre sí. Existe consenso entre los investigadores en que una medida adecuada de la carga de enfermedad debe referirse al tiempo. Aunque más adelante abundaremos en estos conceptos, señalamos aquí que se han desarrollado una serie de medidas basadas en el tiempo para evaluar la carga de enfermedad. En concreto la recomendada por la OMS es el año de vida ajustado por discapacidad (AVAD o Disability- Adjusted Life Year, DALY). El AVAD expresa los años de vida perdidos por muerte prematura y los años vividos con una discapacidad de gravedad y duración específicas. Un AVAD es, por tanto, un año de vida saludable perdido, y combina los años de vida perdidos (AVP), debido a la mortalidad y los años vividos con discapacidad (AVD) en una sola medida. (97).

La suma de AVAD perdidos en una población puede considerarse como la brecha entre el estado de salud actual de la población y una situación ideal en la que toda la población vive hasta una edad avanzada, libre de enfermedades.

Uno de los condicionantes para conocer la carga de enfermedad de determinada condición clínica es tener bien definida su prevalencia. La artrosis constituye una de las principales causas de discapacidad entre la población anciana de nuestro país y de otros países desarrollados; pero a pesar de su relevancia se han realizado pocos estudios de prevalencia. Algunos resultados publicados han sido aproximaciones poco precisas de esta enfermedad, como los extraídos de encuestas de salud tras preguntar por la presencia de trastornos crónicos en general (98).

En general, se estima que la prevalencia global de artrosis en el mundo es muy alta, con cifras de prevalencia superiores a los 300 millones de población afectada en el año 2017, y con un incremento del 31,4% desde el año 2007 (99).

INTRODUCCIÓN

En Estados Unidos se estima una prevalencia general de artrosis para la población de adultos no institucionalizados del 10,5% (25,6 millones) en 2015. La prevalencia ponderada en las mujeres fue del 13,3% en relación con el 7,5% en los hombres; y los blancos no hispanos tuvieron una prevalencia más alta (13,3%) que los afroamericanos (7,5%) o los hispanos (4,2%) (100).

En España, se han publicado 3 estudios transversales en los que se ha calculado la prevalencia de la artrosis. El primero publicado en 1982, estima la prevalencia de artrosis en el 26,1%, con un 20% de la población seleccionada de más de 60 años. El segundo publicado en 1993, donde se diagnostica de artrosis al 23,8% de la muestra, que incluye personas mayores de 18 años (98).

En estudios más recientes, el estudio EPISER 2016 (101), promovido por la Sociedad Española de Reumatología, estima una prevalencia de artrosis del 29,35% (IC 95%: 27,77-30,97) que se incrementa con la edad, alcanzando su valor más alto en los mayores de 80 años (52,6%; IC 95%: 46,97-58,29). La enfermedad era más frecuente en las mujeres, sobre todo a partir de los 60 años. Los datos de prevalencia se corresponden para la artrosis sintomática con confirmación radiológica de cinco localizaciones (columna cervical, lumbar, cadera, rodilla y mano), y a partir de una muestra formada por 3.336 individuos de 40 o más años.

La carga de la enfermedad de la artrosis viene determinada no solo por su prevalencia, sino también con los costes de la enfermedad para las personas y los sistemas de atención médica. En este contexto, los estudios del coste de la enfermedad han revelado el gran impacto económico de la artrosis en el mundo.

Los costes directos incluyen las visitas al médico, hospitalizaciones, pruebas diagnósticas y los costes relacionados con los tratamientos médicos para el control del dolor, así como los relacionados con la cirugía de reemplazo articular. Además, incluye costes no médicos derivados de la ayuda en el trabajo, el hogar y el autocuidado, ayudas para la adaptación, dispositivos, equipo de asistencia para el hogar y transporte.

Los costes indirectos incluyen todo tipo de pérdida de productividad, asociada a bajas temporales o de larga duración, incapacidad laboral permanente. Así como, costes de ayuda con las tareas del hogar, si el paciente no tiene trabajo remunerado fuera de casa (102).

Desde el punto de vista del impacto de la pérdida de funciones, los trastornos musculoesqueléticos ocupan tres de las 25 causas principales de discapacidad tanto en 2005 como en 2015, según el estudio de análisis sistemático de las cargas de enfermedad a nivel mundial

(103). Y en lo que respecta a la artrosis supone la decimotercera causa más común de discapacidad en todo el mundo.

Esta situación está siendo favorecida por el hecho de que se han logrado avances sustanciales en la reducción de la mortalidad, con una ampliación de la esperanza de vida en todo el mundo; lo que ocasiona un cambio epidemiológico, caracterizado por un incremento de las enfermedades no mortales, como es el caso de la artrosis.

Como comentábamos anteriormente los trastornos musculoesqueléticos constituyen una de las principales causas de discapacidad o dicho de otra forma de años de vida con discapacidad (AVD). Se estima un incremento del 34,8% en los años de vida con discapacidad (AVD) entre 2005 y 2015, e incluso un incremento de las tasas estandarizadas por edad, en los años de vida con discapacidad (AVD) del 3,90% (3,00-4,83%) en 2015, a diferencia del descenso que se observa en la mayor parte de las enfermedades no transmisibles (104).

Las principales causas de discapacidad varían considerablemente con la edad. La artrosis constituye una de las principales causas de discapacidad en los adultos de mediana edad y por encima de los 65 años (104).

La carga global de enfermedad, definida como se ha hecho previamente, mensura el impacto de una condición clínica en el conjunto de la sociedad, pero hay determinados matices sobre el impacto que recibe cada uno de los individuos que no queda caracterizado nítidamente por esta medida. La afectación sobre su salud percibida por el individuo, la carga de enfermedad “personal” puede tener otras manifestaciones aparte de la discapacidad, más intangibles, como pueda ser la pérdida de bienestar. Esta pérdida también puede ser capturada a través de ciertas medidas del estado de salud como la calidad de vida relacionada con la salud (CVRS) (105), como explicaremos más adelante.

La medición de la CVRS ha proporcionado herramientas para evaluar el impacto de determinada situación clínica en un sujeto, una de las cuáles es la utilidad, que se definirá más precisamente en capítulos posteriores. La utilidad es una forma de medir la preferencia de un sujeto por un estado de salud determinado y se suele clasificar en una escala donde el 0 se identifica con el peor estado de salud posible, la muerte, y el 1 con un estado de salud perfecta (106). En este contexto, la pérdida de utilidad atribuida a la artrosis en los estudios, generalmente ha sido de aproximadamente 0,10 puntos cuando se ajusta a otras variables que podrían afectar a la Calidad

INTRODUCCIÓN

de Vida Relacionada con la Salud (CVRS) percibida, una pérdida aproximadamente tres veces mayor que la producida por otras afecciones crónicas, como la diabetes o el asma (107).

La discapacidad en los pacientes con artrosis parece relacionada con la limitación funcional, y la interferencia asociada al dolor para realizar actividades. Los pacientes con artrosis experimentan un significativo incremento del grado de limitación para realizar actividades de la vida diaria por el dolor, así como mayores limitaciones funcionales, en comparación con los pacientes sin artrosis. Se estimó que las interferencias para realizar actividades directamente relacionadas con el dolor duplicaban a las de la población sin artrosis (100).

También se ha descrito una probabilidad significativamente mayor de limitaciones cognitivas. Se puede intuir que estas limitaciones funcionales resultantes del dolor tengan un impacto sustancial en la productividad, el impacto individual de la enfermedad y la carga global de enfermedad. Las cuales están directamente relacionadas, aunque puedan evaluarse con diferentes herramientas que ofrecen una visión global del problema.

2.3. La artrosis de miembro inferior

El término OAMI, engloba la afectación en forma de artrosis de las articulaciones de la cadera, rodilla, tobillo, junto a las asociadas a los huesos del tarso y las pequeñas articulaciones de las falanges y metatarsfalángeas.

La cadera y la rodilla son las principales articulaciones sinoviales afectadas por la artrosis, con una prevalencia muy superior al resto, así como una mayor limitación funcional, lo que globalmente se traduce en una de las principales causas de demanda en la atención médica en lo referente a la patología del aparato locomotor.

Es por ello que nos vamos a centrar dentro de la artrosis de miembro inferior en la localizada en la cadera y la rodilla.

Si bien la artrosis de cadera y rodilla aumentan con la edad (15), como desarrollamos al hablar de los factores de riesgo, la evidencia sugiere que la artrosis en estos dos sitios articulares tiene diferentes determinantes. En particular, la artrosis de rodilla tiene una asociación más fuerte con la presencia de artrosis en otras articulaciones como la mano, lo que sugiere una predisposición sistémica más generalizada a la enfermedad.

Algunas de estas diferencias de comportamiento, junto con el hecho de que la artrosis de cadera parece tener un componente más hereditario que la artrosis de rodilla (108); y que los estudios genéticos indican poca correlación genética entre los dos trastornos (109), han respaldado la idea de considerar la artrosis de cadera y rodilla como enfermedades diferentes desde un punto de vista etiopatogénico, aunque no así desde la perspectiva de la práctica clínica en la actualidad.

Desde el punto de vista clínico, la artrosis de rodilla se caracteriza por dolor de características mecánicas de inicio gradual, y disminución de la movilidad articular. Muchas personas con artrosis experimentan dolor y dificultades para andar, cambiar de postura de sentada a levantada y, sobre todo, para subir o bajar escaleras. Es frecuente que las personas afectadas requieran del uso de ayudas externas para la deambulación y precisen subir las escaleras una a una evitando la carga en la rodilla más sintomática (110).

Estos síntomas se asocian con frecuencia a una sensación de inestabilidad o fallo de la rodilla. La sensación de bloqueo de la rodilla puede estar relacionada con la rigidez, la presencia de cuerpos

INTRODUCCIÓN

libres intraarticulares secundarios al desprendimiento de fragmentos condrales o por la presencia de lesiones meniscales asociadas.

Las rodillas con artrosis presentan a menudo crepitación, aunque este síntoma es más indicativo de afectación femorrotuliana que femorotibial; por lo que es valorable en las artrosis femoropatelaes aisladas o en estadios evolucionados de artrosis donde se produce un desgaste completo de la superficie articular evolucionando en una artrosis tricompartmental.

La palpación de la parte interna y/o externa de la interlinea articular puede causar dolor, e incluso es frecuente el dolor a la palpación de la metáfisis tibial proximal por sobrecarga mecánica o tendinitis asociadas de los tendones isquiotibiales en su inserción en la pata de ganso e incluso dolor en el trocánter mayor (111).

El derrame sinovial, cuando está presente, no suele aumentar la temperatura local, ni causar eritema a diferencia de la artritis infecciosa. Si es voluminoso, puede estar asociado a aumento de volumen de la bursa gastrocnemio-semimembranosa, generando un quiste poplíteo o quiste de Baker.

En ocasiones es evidente la desalineación de la extremidad inferior, habitualmente en varo, que es frecuente en personas con artrosis avanzada, pero que puede estar presente incluso en personas con artrosis leve o incipiente.

En cuanto a la artrosis de cadera puede manifestarse como dolor en la región inguinal que empeora con el movimiento (de características mecánicas), diferenciándose de esta forma de otras causas de dolor en la región inguinal, como son las patologías genitourinarias o gastrointestinales.

El dolor de cadera referido por debajo de la rodilla es común con una articulación de la cadera degenerada y sigue la distribución del nervio safeno, que se ramifica desde el nervio femoral. También es frecuente el dolor en la cara anterior del muslo, o la nalga y la región trocantérea (112).

En estos casos se hace necesario establecer el diagnóstico diferencial con otras patologías, como la meralgia parestésica, es decir, la compresión del nervio cutáneo femoral externo; bursitis trocantérea, patología degenerativa de la columna lumbar o la claudicación de miembros inferiores por causa vascular. También es importante descartar otras causas de dolor inguinal de

origen traumatológico como la osteonecrosis de la cabeza femoral, fracturas por estrés del cuello femoral o ramas pélvicas, lesiones musculares (síndrome del psoas iliaco, síndrome del piriforme, tendinitis de los aductores) etcétera.

Las personas con artrosis de cadera tienen limitaciones para andar, agacharse y hacer las transferencias, así como para subir escaleras. La rotación interna de la cadera afectada disminuye con frecuencia y puede ser muy dolorosa, incluso en fases iniciales a menudo es evidente en actividades como ponerse los calcetines, atarse los cordones, cortarse las uñas o mantener relaciones sexuales (110).

La deformidad visible, la contractura en flexión de la cadera o una disminución considerable de la movilidad son indicadores de enfermedad más avanzada. Así como el acortamiento visible de la extremidad afecta que es poco frecuente y se limita a casos de artrosis secundaria secuelas de necrosis avasculares, displasia del desarrollo de la cadera o enfermedad de Perthes, entre otras causas.

Como mencionamos al hablar de la artrosis de forma general, es característica la existencia de una disociación clínico-radiológica, es decir, entre la descripción de síntomas clínicos referidos por el paciente y la existencia de alteraciones anatómicas presentes en los estudios radiológicos.

Desde este punto de vista podemos hablar que para establecer el diagnóstico de artrosis de cadera o rodilla podemos utilizar criterios clínicos, radiológicos o una combinación de ambos, que constituye la forma más extendida de diagnóstico en el ámbito de la salud pública.

Por lo tanto, el diagnóstico de artrosis de miembro inferior se puede realizar usando la evidencia clínica, por la presencia de tres de los seis signos y síntomas más útiles en el diagnóstico, como son el dolor persistente, movimiento restringido, rigidez matutina, crepitación, agrandamiento óseo y función articular reducida (113) o basándonos en la evidencia radiológica (79).

Un estudio reciente ha demostrado que en articulaciones particulares como la cadera, puede no llegarse al diagnosticarse la artrosis, si el diagnóstico se basa únicamente en radiografías (114). Por lo tanto, la recomendación basada en la evidencia para el diagnóstico de artrosis de rodilla y cadera debe incluir características tanto radiográficas como clínicas; de acuerdo con los criterios radiológicos y clínicos del American College of Rheumatology (ACR) para la artrosis de rodilla (115) y cadera (95).

INTRODUCCIÓN

En el caso de la artrosis de rodilla, la combinación de los criterios con mayor sensibilidad y especificidad incluyen el dolor de rodilla (la mayoría de los días del mes anterior) y la presencia de osteófitos con alguno de los siguientes: líquido sinovial característico de artrosis, edad superior o igual a 50 años, rigidez matutina de menos de 30 min y crepitaciones con la movilidad activa de la rodilla. Para el diagnóstico es necesario la combinación de los dos primeros criterios o la presencia de dolor con al menos 3 del resto de los criterios.

En el diagnóstico de artrosis de cadera es necesaria la presencia de dolor en la cadera durante la mayoría de los días del mes anterior, combinado con al menos dos de los siguientes criterios, la presencia de osteofitos radiográficos femorales y/o acetabulares, pinzamiento radiográfico del espacio articular de la cadera o una velocidad de sedimentación globular (VSG) menor o igual a 20.

Para la cadera, la sensibilidad y la especificidad de los criterios del American College of Rheumatology (ACR) son del 89% y del 91% (95), respectivamente, frente al 91% y al 86% para la rodilla (94). Debido a su alta especificidad, estos criterios son más útiles para distinguir la artrosis de la artropatía inflamatoria, pero son menos útiles para distinguir a las personas con artrosis incipiente de las personas sanas.

Los estudios de investigación clínica se apoyan en la valoración radiológica para definir la presencia y gravedad de la artrosis. La radiografía simple es una prueba de fácil y rápido acceso en el sistema sanitario actual, lo que la convierte en una prueba muy útil para confirmar el diagnóstico clínico de artrosis de miembros inferiores.

El sistema de gradación semicuantitativo de Kellgren y Lawrence sigue siendo el más utilizado para establecer el diagnóstico radiológico de la artrosis de miembro inferior (93). Esta escala contempla los signos clásicos radiológicos como son el estrechamiento de la interlinea articular, la presencia de osteofitos, esclerosis subcondral y la deformidad de los extremos óseos. Comprende cinco clases de 0 a 4 que corresponden a ausencia de artrosis, artrosis dudosa, leve, moderada y grave.

En el grado 0 (ausencia de artrosis) no encontramos osteofitos, ni pinzamiento del espacio articular. En el grado 1 se observa dudoso estrechamiento del espacio articular, con posible labio osteofítico. El grado 2, que establece, por lo general, diagnóstico de artrosis, está caracterizado por la presencia de osteofitos con posible estrechamiento del espacio articular. El grado 3 viene

definido por la presencia de osteofitos moderados y estrechamiento evidente del espacio articular, con algo de esclerosis y posible deformidad del extremo óseo. Finalmente, el grado 4 (artrosis grave) está caracterizado por la presencia de osteofitos grandes, estrechamiento marcado del espacio articular, con esclerosis grave y deformidad definida de los extremos óseos.

El sistema de Kellgren y Lawrence tiene varias limitaciones, ocasionados por la variabilidad de las definiciones que encontramos en la bibliografía, con la existencia de varias versiones (116), y por la categorización de todos los estrechamientos o pinzamientos “evidentes” del espacio articular como grado 3 (artrosis moderada), lo que disminuye la sensibilidad a los cambios. Para ello se han propuesto modificaciones en la clasificación que incorporan el uso de la resonancia magnética, que nos permite obtener imágenes de la morfología y composición del cartílago, dado que el pinzamiento del espacio articular no siempre es un hallazgo específico de pérdida del cartílago articular, si no que puede ser consecuencia de una lesión meniscal (117).

En general, con una buena anamnesis del cuadro clínico y con los hallazgos del estudio radiológico, pocas veces está indicado realizar pruebas analíticas para establecer el diagnóstico de artrosis de rodilla y/o cadera. La determinación del factor reumatoide o los anticuerpos antinucleares debe reservarse para los pacientes con hallazgos que sugieren una artropatía inflamatoria.

Respecto al estudio del líquido sinovial, que constituye uno de los criterios menores de la American College of Rheumatology (ACR) para el diagnóstico de artrosis de rodilla. Suele ser normal o levemente inflamatorio, de incoloro a ligeramente amarillento, con una cifra de leucocitos menor o igual a 2.000 células/mm³ (< 2 células por 10 campos de gran aumento) (118).

Es fundamental poder definir con criterios objetivos y uniformes el diagnóstico de artrosis para llegar a conocer el verdadero impacto de la artrosis de miembros inferiores en la población. Los estudios que encontramos en la bibliografía emplean una diversidad de criterios diagnósticos de artrosis, aunque la artrosis de rodilla es la que tiene los criterios mejor definidos y es más fácil de valorar.

Las prevalencias son globalmente mayores cuando se utiliza solo la evidencia radiológica, seguidas de las observadas al definir la enfermedad por los síntomas y son menores cuando se combinan ambos criterios (98).

Dada la progresión de la artrosis de miembros inferiores con la edad como factor de riesgo, es lógico pensar que las mayores prevalencias se obtengan cuando se utilizan exclusivamente

INTRODUCCIÓN

criterios radiológicos para establecer el diagnóstico, pues se detectan casos de artrosis incipiente radiológica que aún no presentan síntomas. De hecho, ningún estudio se basa exclusivamente en la evidencia radiológica para estimar impacto de la artrosis de miembro inferiores, ya que lo más relevante en estos casos, son los síntomas.

Puede resultar contradictorio que se describe una disminución en la prevalencia de artrosis de miembros inferiores en edades muy avanzadas (mayores de 80 años) en los estudios que incluyeron criterios sintomáticos, ya sea exclusivamente o combinados con criterios radiológicos. Pero este hecho podría explicarse por las características de las personas que alcanzan esta edad que tendrían mejores condiciones generales de salud, y sobre todo, una mayor tolerancia y adaptación al dolor y la discapacidad en comparación a las personas de menor edad (98).

2.4. Carga de enfermedad de la artrosis de miembro inferior

Como referimos anteriormente, la carga global de enfermedad, es una medida que mensura además del impacto sobre la sociedad de la muerte prematura, aquel otro producido por la discapacidad que conlleva (97).

Para una determinada enfermedad, su carga global de enfermedad, viene condicionada fundamentalmente por su prevalencia y por la discapacidad que produce.

En cuanto a la prevalencia, la artrosis de rodilla es la de mayor prevalencia en el miembro inferior con diferencia. Se estimó una prevalencia global de artrosis sintomática de rodilla confirmada radiográficamente en 2010 de 3,8% (IC 95%: 3,6 – 4,1%), siendo mayor en mujeres (4,8% de media) que en los hombres (2,8% de media). No existe evidencia de un cambio en la prevalencia ajustado por la edad estandarizada entre 1990 y 2010 tanto para hombres como para mujeres; con un punto máximo alrededor de los 50 años (119).

La prevalencia de artrosis de rodilla en 2010 fue más alta en los países con ingresos elevados de Asia Pacífico, seguida de Oceanía, África del Norte y Oriente Medio.

La artrosis de cadera es menos común que la de rodilla, se estimó una prevalencia global confirmada radiográficamente en 2010 de 0,85% (IC 95%. 0,74 – 1,02%). Al igual que ocurre con la artrosis de rodilla no existe evidencia de incrementos en la prevalencia estandarizada por edad entre 1990 y 2010 ni para hombres, ni para mujeres. Si bien la prevalencia aumentó constantemente con la edad, a diferencia de la gonartrosis donde existe un fuerte aumento de la incidencia entre los 50-75 años, que es más limitada después de los 75 años (120).

La prevalencia de coxartrosis en 2010 fue más alta en las regiones con ingresos altos de América del Norte seguida por América del Sur y países con ingresos elevados de Asia Pacífico.

Aunque la prevalencia ajustada por edad no parecía haber aumentado hasta 2010, dado que estamos asistiendo a un incremento de la esperanza de vida, y la prevalencia de la artrosis en miembro inferior se incrementa de forma global con la edad, se prevé un incremento de la prevalencia en la población general del mundo.

INTRODUCCIÓN

En 2015, según el estudio Global Burden of Disease, el número de casos de gonartrosis sintomática en el mundo se estimaba en casi 200 millones, un aumento de casi el 33% con relación a 2005 (121). Cifras similares a las documentadas para los casos de coxartrosis sintomática que se estimaba en casi 36 millones, con un aumento del 33,5% para el mismo periodo de tiempo.

En los últimos 20 años se ha observado un aumento de la prevalencia de la gonartrosis sintomática. Este aumento era de alrededor del 10% en los varones y del 6% en las mujeres de la cohorte Framingham (122).

En cuanto a la prevalencia de artrosis de miembro inferior en España. La prevalencia de la artrosis de rodilla fue 13,83% (IC 95%:12,66-15,11) siguiendo criterios de diagnóstico clínico-radiográfico en más del 90% de los casos. Siendo mayor en las regiones del Mediterráneo con una asociación cercana a la significación estadística. La prevalencia de artrosis de cadera en España se estimó en 5,13% (IC 95%: 4,40-5,99), siguiendo los mismos criterios diagnósticos (101).

Si comparamos los datos de prevalencia con los publicados en otros países, vemos que la prevalencia de artrosis de cadera es similar a los datos publicados en poblaciones de EE. UU., Portugal, Holanda y Francia. Sin embargo, la prevalencia de la artrosis sintomática de rodilla es superior a la referenciada en EE. UU., Francia y Noruega, y más parecida a la publicada en Portugal, Canadá y Holanda.

Los costes sanitarios por artrosis de miembro inferior en Europa, se estiman en aproximadamente 5.000 euros por año, y 12.000 euros por año en los EE.UU (123). Esta diferencia tan sustancial se puede explicar por la mayor frecuencia de obesidad y comorbilidades relacionadas en Estados Unidos que en países europeos (124).

Se espera que la carga para la sociedad sea alta debido a la elevada prevalencia de enfermedad sintomática, con unos costos inducidos que podrían ascender entre 350 y 700.000 millones de euros / año de costes directos en Europa (123).

En España, se ha informado de un coste anual de 1.500 euros en 2007 para pacientes con artrosis de rodilla o cadera, de los cuales el 86% fueron costes directos. Los costes indirectos, aquellos relacionados con las pérdidas de productividad, representan en torno al 14%, y tienen su componente más importante al relacionado con la ayuda a las amas de casa en el hogar (125). En cifras absolutas se estimó un coste nacional de 4.738 millones de euros, lo que representa el 0,5% del producto nacional bruto.

Los costes anuales totales pueden oscilar entre 1.400 € a 1.600 € por paciente, donde los costes más elevados se asocian a pacientes con mayor comorbilidad y con grados de artrosis moderada o grave (grados 3 y 4) según la escala radiológica de Kellgren y Lawrence en comparación con los pacientes con artrosis dudosa o leve (grados 1 o 2). Los pacientes con edades iguales o superiores a los 60 años también generan costes médicos directos más altos en comparación con el grupo de edad de 50 a 59 años (125).

Desde el punto de vista del impacto de la pérdida de funciones, la artrosis supone la decimotercera causa más común de discapacidad en todo el mundo (121), y concretamente la artrosis de rodilla y cadera ocupa el undécimo puesto, medido como años de vida con discapacidad (AVD) en el año 2010 (126).

Al igual que ocurre al hablar de la artrosis en general, se estimó un incremento en los años de vida con discapacidad (AVD), que para la artrosis de rodilla y cadera es del 22% entre 1990 y 2010. Los años de vida ajustados por discapacidad (AVAD) en todas las edades para la artrosis de cadera y rodilla aumentaron de 0,42% del total en 1990 a 0,69% del total de AVAD en 2010 (127).

Por otra parte, el aumento de la esperanza de vida en la mayoría de los países, ha planteado el debate en algunos países con índices sociodemográficos altos, de la ampliación de las edades de jubilación para reflejar estos cambios en la supervivencia y asegurar los sistemas de protección social, especialmente necesarios en las últimas etapas vitales.

Un factor crucial en estos debates es si el aumento en la esperanza de vida, va acompañado o no, del incremento de los niveles promedio de salud funcional. Encontramos que los AVAD estandarizados por edad per cápita disminuyeron (no en el caso de la artrosis), pero bastante lentamente, durante el período 2005 a 2015 (121).

En este aspecto, la artrosis, constituye una de las principales causas de discapacidad en la población de mediana edad en los países desarrollados, con unos índices de años de vida con discapacidad en ligero aumento, pero con los años de vida perdidos (AVP) disminuyendo, y lo que es más significativo todavía con unas cifras de prevalencia incrementadas en un 33% en 10 años en el caso concreto de la artrosis de miembro inferior. Por todos estos datos es de esperar un incremento sustancial de las bajas laborales y jubilaciones anticipadas consecuencia de la artrosis de miembro inferior.

INTRODUCCIÓN

El diagnóstico clínico de artrosis de cadera o de rodilla se asoció con dificultades informadas por el paciente en la movilidad, la capacidad de autocuidado y la realización de actividades habituales (128).

La asociación entre la artrosis clínica de la rodilla y las dificultades con la movilidad, el autocuidado y la realización de las actividades habituales se mantuvo estable después del ajuste de los factores de confusión, pero para la artrosis clínica de la cadera solo la asociación con los problemas para realizar las actividades habituales siguió siendo significativa después del ajuste correspondiente (128)

Además, encontramos una pérdida de productividad entre los adultos con artrosis que está mediada directamente por la presencia de dolor e indirectamente por el impacto del dolor en la función física.

La discapacidad asociada a la artrosis tiene que ver con dos características clínicas fundamentales, el dolor y la pérdida de función articular (100).

El dolor juega un papel clave en el riesgo del abandono del mercado laboral. Se observó una fuerte asociación entre la interferencia del dolor para realizar actividades y la pérdida del trabajo, especialmente en la población con artrosis de rodilla (OR: 1,52; IC del 95%: 1,16–1,99; $p = 0,002$), mientras que no se observó una asociación significativa para la artrosis de cadera (OR: 1,04; IC 95%: 0,36–2,99; $p = 0,938$) (129).

De manera similar a lo que se observó con el dolor, la asociación entre artrosis de rodilla y salida del trabajo se vuelve no significativa si solo se analiza el subconjunto de población con menos discapacidad. Encontramos una relación casi lineal entre los niveles de discapacidad y la probabilidad de salida del mercado laboral (129).

Si nos referimos al impacto de la enfermedad a nivel individual, se sabe que la artrosis es una enfermedad crónica que tiene un gran impacto en la Calidad de Vida Relacionada con la Salud (CVRS). Los adultos con artrosis tienen peor estado de salud físico y mental percibido, comparado con los adultos sin artrosis. Encontramos un mayor impacto en los componentes físicos de la CVRS en relación con los componentes mentales, lo que es de esperar en una enfermedad que afecta a las articulaciones y se caracteriza por una función reducida (100).

En España, cuando se utilizó el cuestionario de salud EQ-5D-3L para evaluar la CVRS en un estudio de artrosis, la identificó como una afección crónica que disminuye la utilidad en la misma medida que los problemas cardiacos o la diabetes mellitus, con una disminución en la utilidad de hasta 0,10 puntos en comparación con la población general mayor de 65 años (130).

Si nos centramos en la artrosis de miembro inferior, la artrosis de cadera y rodilla tiene un impacto muy significativo en cuanto a merma de la calidad de vida. Vivir con artrosis de rodilla o cadera implica una pérdida de utilidad de aproximadamente 0,30 en promedio en comparación con la población general sin artrosis, que excede los umbrales de la relevancia clínica obtenida en España (0,07 puntos en el índice de utilidad EQ-5D-5L) (131).

La presencia de artrosis de cadera disminuye la utilidad en un promedio de 0,347 puntos para las personas que son similares en términos de edad, sexo, grupo social, estado de convivencia y carga de enfermedad crónica. Para los pacientes con artrosis de rodilla, la disminución de este valor es de 0,295 puntos en promedio.

El gran impacto de la artrosis de miembros inferiores no solo abarca los dominios del dolor y la pérdida de función, que son más evidentes para el clínico, si no que los datos muestran un deterioro del componente psicológico que aparece incluido en el dominio de la ansiedad/depresión del cuestionario EQ-5D-5L (131). Este impacto de la artrosis de miembro inferior en la salud mental también ha sido descrito por otros autores; concretamente se determina que las personas con artrosis de rodilla o cadera tienen mayor probabilidad de desarrollar síntomas depresivos en comparación con las personas sin artrosis (132).

2.5. Medición de los resultados en salud desde la perspectiva del paciente

La evaluación de la salud, y por ende, de las intervenciones que se hacen para mejorarla, requiere de la definición de este constructo¹. No es fácil acotar el concepto de salud, pues el hecho de que todas las personas tengan un cierto grado de salud, y que ésta adquiera numerosas formas, tanto en personas diferentes, como en diferentes momentos de la vida de la misma persona, indican que todas esas formas podrían parecer aproximaciones parciales a un concepto único (133).

No existe una única definición de salud, y dicha definición ha sufrido una evolución histórica que se ha desarrollado de forma paralela a la concepción de la enfermedad, y que se ha acompañado de un cambio de las herramientas para evaluarla.

La salud es un componente fundamental del bienestar en las sociedades desarrolladas. La acepción de salud ha ido cambiando a lo largo de los años desde un concepto meramente biologicista hasta abarcar dimensiones emocionales, sociales o cognitivas.

En el siglo XIX Claudio Bernard introdujo la medicina al mundo de la ciencia, y entre sus reflexiones acerca de la salud que documenta en su libro: "Introducción al estudio de la medicina experimental", establece que la condición para una vida sana no se encuentra en el organismo ni en el medio externo, sino en ambos. Señala que la "relativa constancia" del medio interno, a pesar de las variaciones del medio externo, se traduce en el funcionamiento de los órganos y sistemas corporales dentro de un margen de normalidad. Por eso consideraba que la salud del hombre depende de una "relación necesaria de intercambio y equilibrio" al interactuar con fuerzas físicas, objetos materiales, sustancias químicas y seres biológicos del ambiente externo (134).

Perkins, en 1938 basado en los conceptos de Bernard define la salud como "un estado de relativo equilibrio de la forma y función corporal, que resulta del ajuste dinámico del organismo ante las fuerzas que tienden a alterarlo. Y añadía, "no es, pues, el resultado de la interrelación pasiva entre las sustancias del organismo y los factores que pretenden romper la armonía con el medio

¹ Un "constructo" se define como una construcción teórica para comprender un problema determinado.

externo, sino la respuesta activa de las fuerzas corporales que funcionan para establecer los ajustes para preservar la vida" (134).

En 1941, Sigerist fue el primero en diferenciar salud de enfermedad, al definir que la medicina debía desarrollar cuatro grandes tareas: la promoción de la salud, la prevención de la enfermedad, el tratamiento del enfermo y la rehabilitación. Fue el primero en asumir que "la salud no es simplemente la ausencia de enfermedad, es algo positivo, una actitud gozosa y una aceptación alegre de responsabilidades que la vida impone al individuo"(135).

El concepto moderno de salud se hace presente cuando la Organización Mundial de la Salud (OMS), en el momento de su creación en 1948, adopta en su Constitución la definición de salud propuesta por Andrija Stampar, uno de los fundadores de la OMS (136). En esta definición, globalmente aceptada, se contempla la salud como "un estado de completo bienestar físico, mental y social, y no sólo la ausencia de afecciones o enfermedades" (137).

Esta definición generó numerosas críticas, fundamentalmente por tratarse de una concepción estática, y por equiparar bienestar a salud.

Una de las objeciones más comunes a la definición de salud de la OMS es su categorización de la salud como estado. Es una noción estática. Los que critican el uso de este término tienden a ver la salud más como un "proceso" o una condición dinámica que implica ajustes continuos a las demandas cambiantes del entorno físico y social (138). En 1960, Sigerist empieza a valorar la salud desde una perspectiva dinámica cuando afirma: "vivimos un ritmo específico, determinado por la naturaleza, la cultura y los hábitos. Un ritmo inalterado es la salud, la enfermedad es la interrupción de éste" (139).

En cuanto al carácter utópico de la definición de salud de la OMS, las críticas no se hicieron esperar. René Dubos en 1956 escribe que "la salud es un espejismo que merece la pena buscar, aunque es esencialmente inalcanzable". Define la salud como un estado físico y mental razonablemente libre de incomodidad y dolor, que permite a la persona en cuestión funcionar eficazmente por el tiempo más largo posible en el ambiente donde por elección está ubicado" (134).

Otras muchas críticas se han sumado a la definición de salud de la OMS, como la falta de consenso en el significado de bienestar, o sus diferentes acepciones en distintas culturas, o su carácter "eufórico" que catalogaría a la mayor parte de la población como enfermos (140). Una derivada

INTRODUCCIÓN

fundamental de estas críticas es el limitado papel de la medicina para conseguir mantener o mejorar la salud bajo esta definición. El propio Dubos puso en cuestión el papel de la medicina en la consecución del bienestar: “la medicina no puede determinar por sí misma la calidad de vida; sólo puede ayudar a las personas a alcanzar un estado de salud que les permita cultivar el oído de la vida, pero a su manera. Esto implica la posibilidad de disfrutar de las satisfacciones fundamentales de la vida biológica y la posibilidad de que cada persona haga lo que quiera y se convierta en lo que quiera, de acuerdo con los valores humanos que trascienden el juicio médico”(141).

Con el fin de corregir parte de los defectos de la definición de Stampar que fue aceptada por la Organización Mundial de la Salud (OMS), Milton Terris sugiere distintos grados de salud como los hay de enfermedad. Afirma que la salud puede ser evaluada bajo dos prismas: uno subjetivo y otro objetivo. El aspecto subjetivo se refiere a la sensación de bienestar, el aspecto objetivo a la capacidad de funcionamiento. Considera que hay que suprimir la palabra “completa” de la definición de salud, pues la salud no es absoluta, sino que hay grados de salud. Además propone sustituir el término “enfermedad” (illness en inglés) por "dolencia" (disease en inglés), ya que la salud y la dolencia no son mutuamente excluyentes. Apuesta por una definición de la salud tanto en términos funcionales como subjetivos con la siguiente expresión: "La salud es un estado de bienestar físico, mental y social y de capacidad de funcionamiento, y no sólo la ausencia de enfermedad o dolencia" (142). Como puede verse esta definición elimina el término “completo estado de bienestar”, que convertía a la salud en un concepto utópico. Por otra parte mantiene el componente subjetivo del bienestar, pero añade un elemento de gran importancia como es la capacidad de funcionamiento de la persona, introduciendo de esta forma la posibilidad de medir el grado de salud.

Así desde principios de los años 70 del pasado siglo se empiezan a reconocer dos tipos de aproximaciones a la definición de salud: una definición idealista en la que la salud está abierta, que se identificaba con las propuestas de Milton Terris y Dubos, y se englobaban en un modelo de “bienestar”, y otra más elástica en el que la salud se relacionaba con las tensiones y la interacción con el entorno que se denominó modelo “ambiental”. El modelo “ambiental” de salud, cuyos elementos se veían ya presentes en las definiciones de salud de Claude Bernard primero, y de Perkins o de Sigerist después, propone que la vida y la salud son en gran medida

una cuestión de adaptación a nuestro entorno. Si nos adaptamos bien, el estrés y la enfermedad resultante se minimizan (140).

Dado que el objetivo de este trabajo no es el acotar definitivamente el concepto de la salud, concluimos, como hacen otros autores, aceptando que la definición de la salud debe ser un esfuerzo continuo, que implica a diversos actores, incluidos los pacientes. Quizá no se deba aspirar a generar un acuerdo sobre el significado preciso de la salud, pero debe crearse un marco que permita identificar y medir los complejos componentes y determinantes de la salud (138).

Sobre lo que sí hay acuerdo, es sobre el hecho de que las definiciones simplistas de la salud conducen a medidas de la salud, o de los resultados sanitarios y de la calidad de la atención que también son simplistas (140). El interés sobre cómo medir la salud está relacionado con la evaluación de la atención para la salud, y se elige deliberadamente para reflejar los problemas de interés social o valores fundamentales. La elección de determinadas medidas de salud son acciones que reflejan y orientan los objetivos sociales y políticos.

La evaluación de los estados de salud, y por ende de los resultados de las intervenciones en salud, se ha basado tradicionalmente en la medición de variables "objetivas" que identificaban dicho estado con determinadas características mensurables mediante herramientas o tecnologías de laboratorio. Si se valora de una manera histórica el acercamiento al paradigma de enfermedad y a la medición de resultados se pueden señalar tres grandes etapas, un primer momento de predominio de enfermedades infecciosas, un segundo momento de enfermedades crónico-degenerativas y un tercero el predominio de las enfermedades asociadas con los estilos de vida, podremos encontrar diferentes definiciones de salud y diferentes medidas de resultado (143).

En una primera etapa existía un claro predominio de las enfermedades infecciosas y la definición de enfermedad tomaba una perspectiva orgánica y la medición de la salud se hacía a través de parámetros como la mortalidad y morbilidad. (143). La mortalidad por causas específicas y la mortalidad infantil permiten posicionar las poblaciones a lo largo de la transición epidemiológica. En los países que ya han experimentado esta transición epidemiológica, no es previsible un gran cambio en las tasas de mortalidad a medio plazo, y su utilidad es limitada (144).

Posteriormente, en la primera mitad del siglo pasado comienzan a tomar protagonismo las enfermedades crónico-degenerativas, y la definición del estado de salud adquiere un carácter más "funcional". La causa subyacente de la enfermedad es inespecífica y multifactorial. El objetivo de

INTRODUCCIÓN

la atención médica no es solo curar sino también disminuir el dolor, la molestia y la incapacidad. El estado de salud de la población comienza a definirse mediante variables sobre el uso de los servicios, sobre algún resultado del sistema, sobre incapacidad y alguna medida de duración. Para evaluar la atención se utilizan medidas de estructura y de proceso de los servicios de salud, cuyo uso sólo es válido en el caso de que se haya demostrado su asociación con el resultado esperado (143). Son ejemplos de medidas de resultado utilizados en este caso los registros de hospitalizaciones por diferentes causas o escalas de gravedad de la enfermedad como el índice de Karnofsky desarrollado a finales de los años 40 para la evaluación de pacientes con cáncer (145), o la escala de la NYHA (New York Heart Association) para evaluar la gravedad de los síntomas anginosos.

En la segunda mitad del siglo pasado se manifiestan como más prevalentes enfermedades asociadas con un estilo de vida "moderno". En ese contexto las personas son vistas como participantes activos del proceso social y de la atención a la salud. El marco de referencia de la salud se vuelve más amplio al abarcar características biológicas, sociales, psicológicas, etcétera y nace la definición de la salud de la OMS (137). Se comienza a valorar la perspectiva del consumidor en la definición del estado de salud.

En este último contexto aparecen las medidas subjetivas de las condiciones de salud, con una aproximación más holística al cuidado, donde no sólo se busque combatir la enfermedad sino promover el bienestar. En las definiciones más recientes de la salud, como se señaló en párrafos anteriores se propugnaba que la salud podía ser evaluada bajo dos prismas, uno subjetivo referido a la sensación de bienestar, y otro objetivo relacionado con la capacidad de funcionamiento (142). Aceptar una definición de salud más amplia, implica que las medidas subjetivas aparecen como las más sensibles para la detección de deterioros en la salud, porque son percibidas por el propio paciente. Los avances metodológicos en la medición de los resultados de la enfermedad y el tratamiento desde el punto de vista del paciente han dado lugar a diversos nombres para grupos de herramientas que valoran el estado de salud, o el estado funcional, o la calidad de vida, o la calidad de vida relacionada con la salud (CVRS), cuya característica fundamental es la medición de los componentes objetivos y subjetivos del funcionamiento y el bienestar humanos (146).

La incorporación de la medida de la CVRS como una medida necesaria ha sido una de las mayores innovaciones en las evaluaciones de la salud (147).

La calidad de vida relacionada con la salud (CVRS), es una medida de salud que, al igual que la mortalidad, la morbilidad y los parámetros clínicos, tiene identidad propia. Añade a este conjunto la medición de la percepción (del sujeto) de la enfermedad, la salud y sus consecuencias. De manera formal, puede definirse como “el valor asignado por individuos, grupos de individuos y la sociedad a la duración de la vida modificada por las deficiencias, los estados funcionales, las percepciones y las oportunidades sociales, que están influidas por las enfermedades, las lesiones, el tratamiento (médico) o la política (salud)”(148).

2.6. La calidad de vida relacionada con la salud

La necesidad de medir la calidad de vida relacionada con la salud (CVRS) nació en el campo de la salud de manera “paralela” a la necesidad percibida por otros investigadores en otras áreas (como desde la sociología o la economía) de describir el estado de las poblaciones más allá de indicadores tradicionales como los económicos. La aplicación de nuevas tecnologías, con capacidad de prolongar la vida a cualquier precio, la complicada decisión de cantidad versus calidad de la vida, y el dilema ético de la distribución de los recursos económicos en salud, ponen de relevancia la necesidad de conocer las opiniones de los pacientes (149).

La CVRS aparece como un elemento que hace interesantes aportaciones a la caracterización de la enfermedad, especialmente a aquella con características de cronicidad. Las medidas fisiológicas, aportan información al médico, pero tienen un interés limitado para los pacientes y sus familias, especialmente cuando existe una mala correlación con el bienestar que este percibe. Por otro lado, dichas medidas no son suficientes para explicar la percepción del estado de salud de una persona, siendo frecuente que dos pacientes con criterios clínicos similares, tenga una percepción de su salud muy diferente (147).

Frecuentemente el concepto calidad de vida relacionada con la salud (CVRS) se ha utilizado de forma intercambiable con el de calidad de vida (CV), o con el de percepción del estado de salud, asumiendo que los tres constructos estarían evaluando las mismas dimensiones (105). Sin embargo, existen elementos que pueden ayudar a delimitar estos conceptos.

El término calidad de vida (CV) precedió al uso de la CVRS. La CV adquirió relevancia en la literatura médica en la década de los 60' del siglo pasado a medida que diferentes tratamientos demostraban capacidad de prolongar la duración de la vida, a veces sin mejorar la percepción de bienestar de los pacientes (150). Definir la CV resultó una tarea desafiante, aceptándose que el concepto podía acotarse como "un juicio cognitivo consciente de satisfacción con la propia vida" (151) y "una percepción de los individuos de su posición en la vida en el contexto de la cultura y los sistemas de valores en los que viven y en relación con sus objetivos, expectativas, normas y preocupaciones" (152). Los límites de la definición de ambos términos CV y CVRS suelen depender del motivo de la evaluación, pero existen aspectos de la vida ampliamente valorados que no suelen considerarse como salud, como los recursos económicos, el entorno medioambiental o la

libertad individual. Aunque estos elementos, o su ausencia, pueden afectar negativamente a la salud, estos problemas suelen estar alejados de la capacidad de acción de la intervención sanitaria (147).

Una década después comenzó a evaluarse el estado de salud en un intento de evaluar los resultados de los sistemas sanitarios. Uno de los primeros intentos de medir y valorar la salud fue el "Health Status Index" (HSI). El HSI era una medida genérica de la salud en lugar de una medida específica de la enfermedad o la población, y atribuía preferencias a determinados estados de salud, mediante una escala cardinal basada en juicios de valor (153). También existen diferencias, más sutiles quizás, entre la percepción del CVRS y la percepción del estado de salud. Cuando los pacientes valoran su CVRS hacen más énfasis en aspectos relacionados con la esfera de la salud mental, que cuando se evalúa la percepción del estado de salud, donde tiene un impacto mucho mayor el funcionamiento físico. Estos efectos diferenciales pueden sugerir que los pacientes están evaluando dos constructos diferentes (Smith, Avis, & Assmann, 1999).

Fue la literatura sobre medidas del estado de salud la que introdujo el término CVRS (150). Por ejemplo, se encuentra este término en un manuscrito que valora el concepto de años de vida ajustados por la calidad (AVAC). Según los autores, sería preferible hablar de "año de bienestar" que de AVAC pues aquella denominación se vincula de forma más clara con las condiciones de salud, esto es, denota la calidad de vida relacionada con la salud (154).

La calidad de vida relacionada con la salud (CVRS) surge así como una medida necesaria en el campo de la medicina para evaluar la salud y los cambios como resultado de las intervenciones médicas; que debe limitarse a la experiencia que el paciente tiene de su enfermedad y como esta interfiere en su vida diaria (148).

Existen numerosas definiciones de la calidad de vida relacionada con la salud. Todas se basan en una idea de salud definida como el buen funcionamiento físico, psíquico y social. Algunas enfatizan el valor que tiene para el individuo su percepción de funcionalidad, de sus potenciales deficiencias, y las oportunidades sociales (155). Otras ponen de relieve la medida en que se modifica el valor asignado a la duración de la vida en función de la percepción de limitaciones físicas, psicológicas, sociales y de disminución de oportunidades a causa de la enfermedad, sus secuelas, el tratamiento y/o las políticas de salud (156).

INTRODUCCIÓN

Tras revisar diferentes definiciones de la calidad de vida relacionada con la salud (148), se vuelve a difuminar la diferenciación hecha entre CVRS, CV y de la percepción del estado de salud. Cuando el peso de las definiciones recae sobre el funcionamiento y bienestar, el constructo no parece alejarse mucho de la descripción de la salud, tal como la define la OMS. Cuando las definiciones de la CVRS se centran en los aspectos de la CV que pueden verse afectados por la salud, asemejan ambas definiciones pues es difícil imaginar aspectos importantes de la CV que no se vean afectados, al menos en cierta medida, por la salud. Quizás fuese más preciso centrar el foco de estas definiciones de la CVRS en los aspectos de la CV más afectados por la mala salud (150).

En un intento de normalizar las definiciones de CVRS, la Organización Mundial de la Salud (OMS) creó un grupo de trabajo multicultural de consenso, que definió cuales debían ser las características comunes para caracterizar este constructo. Se consensó que la definición de la CVRS debía contener los siguientes elementos (152):

- Ser subjetivas: Recogen la percepción del paciente.
- Ser multidimensionales: deben revelar diversos aspectos de la vida del individuo, en los niveles físico, emocional, social, etc.
- Incluir sentimientos positivos y negativos.
- Registrar la variabilidad en el tiempo: la etapa vital que se atraviesa y el estadio de la enfermedad.

Aunque este consenso es globalmente aceptado, algunos autores reclaman que la medición de la CV, y por ende de la CVRS, incorpore elementos que permitan alguna medida objetiva (150,157).

No se puede dejar de señalar que los conceptos de CV y de CVRS, son interpretables de manera externa a las normas culturales, patrones de conducta y expectativas de cada persona, teniendo estos aspectos una relevancia especial a la hora de construir y/o adaptar instrumentos para medir la CVRS. Sin embargo es frecuente que las investigaciones de CVRS dejen de lado estos aspectos antropológicos y culturales, asumiendo un sistema único globalizado de valores. Si bien el concepto de CVRS se interpreta referido a la persona con carácter individual, esta percepción está influenciada por, al menos tres marcos de referencia (149):

- Marco histórico: la sociedad adopta parámetros diferentes de calidad de vida en distintos momentos históricos.

Caracterización de la Calidad de Vida Relacionada con la Salud en pacientes con artrosis de miembro inferior y validación de nuevas herramientas para su evaluación.

- Pautas culturales: la percepción de las limitaciones funcionales, el grado de autonomía o la percepción del dolor van a estar condicionados por nuestras pautas culturales, e incluso familiares.
- Pautas sociales, las expectativas que cada individuo tiene en su propia vida tiene una estrecha relación con la clase social a la que pertenece.

La CVRS constituye un constructo complejo, que no se puede medir directamente, por lo cual, es preciso diseñar herramientas que sean capaces de aprehender la percepción del paciente en un conjunto de variables o dominios. Se entiende por dominio o dimensión al área de comportamiento o experiencia que estamos tratando de medir, de manera que su evaluación nos permite valorar aspectos relevantes de la percepción del individuo.

El tener instrumentos que sirvan para evaluar la CVRS ha permitido la utilización de este concepto con diferentes propósitos como aproximarse a la naturaleza, severidad y pronóstico de una enfermedad, evaluar los efectos de un tratamiento, para caracterizar las necesidades de salud de una población, o monitorizar su nivel de salud (158). También se ha utilizado esta medida para informar la toma de decisiones pues se conoce su papel como factor asociado a la utilización de servicios, y como un resultado en sí mismo del uso de esos servicios (159), así como su sensibilidad a diferentes características personales modificables y no modificables (160,161).

Pasamos a revisar en el punto siguiente las herramientas utilizadas para evaluar la Calidad de Vida Relacionada con la Salud.

2.7. Herramientas de medida de la calidad de vida relacionada con la salud

Los instrumentos de medida de la CVRS deben englobar cada una de las dimensiones relevantes que definen la percepción del estado de salud del sujeto. Según el grupo WHOQOL los ítems que conforman los cuestionarios para medir la CVRS se deben agrupar en múltiples dimensiones que engloben el funcionamiento físico (capacidad para realizar diferentes actividades físicas), funcionamiento emocional (sensación de ansiedad, depresión...), funcionamiento social (interferencia del problema en la vida social del individuo), rol funcional (grado en que la salud interfiere en el trabajo y otras actividades diarias), funcionamiento cognitivo, percepción de la salud general y síntomas específicos de la enfermedad (152).

Las herramientas para medir la CVRS se han dividido de diferentes maneras. Se suelen clasificar los instrumentos de medida en genéricos, específicos y un tercer grupo que evalúan las preferencias de los pacientes por determinado estado de salud (146). Aunque esta sea la clasificación a la que nos referiremos posteriormente, se han hecho otras que distinguen instrumentos genéricos, específicos de enfermedades concretas e instrumentos para diversas áreas o dimensiones de funcionamiento (como el dolor, por ejemplo) (148).

Los instrumentos genéricos no hacen referencia, ni están relacionados con ningún tipo de enfermedad concreta. Pueden constituirse como perfiles (varios dominios) o índices (medidas resumen). Estos instrumentos permiten comparaciones entre poblaciones y muestran diferentes aspectos del estado de salud, pero pueden no detectar diferencias individuales y no se focalizan en áreas de interés (162).

Son ejemplos de instrumentos genéricos, el Health Status Index (153), Health Insurance Experiment (163), Nottingham Health Profile (164), Duke Health Profile (165), COOP Charts (166), Medical Outcomes Study (167), EQ-5D 3L (168), SF-36 (169), SF-12 (170), y EQ-5D 5L (171).

Quizá los dos instrumentos genéricos más utilizados en nuestro medio sean el EQ-5D (European Quality of Life-5 Dimensions) y el SF-36 (Short Form – 36).

El EQ-5D es una medida genérica del estado de salud desarrollada en los años 90 del siglo pasado, por un grupo de investigación internacional, el grupo EuroQol. Constituido en el año 1987, el

grupo EuroQol inicialmente comprendía una red de investigadores internacionales, multilingües y multidisciplinarios de siete centros, en Finlandia, los Países Bajos, Noruega, Suecia y el Reino Unido. Actualmente, el grupo está formado por investigadores de Canadá, Dinamarca, Alemania, Grecia, Japón, Nueva Zelanda, Eslovenia, España, Estados Unidos y Zimbabue. El proceso de desarrollo compartido y experimentación dio como resultado el EQ-5D (168), una medida genérica del estado de salud que proporciona un perfil descriptivo simple y un valor de índice único que se puede utilizar en la evaluación clínica y económica de la atención médica y en las encuestas de salud de la población.

El EQ-5D consta por tanto de dos partes: un termómetro para valorar la CVRS de 0 a 100, que varía desde 0 (peor condición de salud) hasta 100 (mejor condición de salud imaginable) según una escala visual analógica (VAS); y un cuestionario formado por 5 preguntas o dimensiones (movilidad, auto-cuidado, realización de actividades habituales, dolor/malestar, y ansiedad/depresión) que primariamente tenía 3 niveles de respuesta (desde sin problemas hasta problemas extremos) que definían $3^5=243$ posibles estados de salud. Aunque el cuestionario había sido ampliamente utilizado, había mostrado algunas limitaciones. El EQ-5D-3L mostraba un efecto de techo (172), por lo que se creó un nuevo desarrollo del mismo cuestionario, denominado EQ-5D-5L en el que cada dimensión se evaluaba con una escala de 5 categorías, (desde 1 sin problemas, hasta 5 problemas imposibilitantes o extremos). Usando combinaciones de estas cinco preguntas se pueden obtener $5^5=3125$ estados de salud, y una puntuación ponderada única, llamada índice de utilidad (concepto que será desarrollado más adelante), de modo que cuanto mayor sea la puntuación, mejor será el estado de salud (171).

La versión EQ-5D-5L en castellano para España ha mostrado buenas propiedades psicométricas, superando las limitaciones del EQ-5D-3L; y por lo tanto podría ser muy útil como medida de resultado, al menos en pacientes con artrosis de cadera o rodilla (173).

El cuestionario de salud SF-36 (Short Form - 36) fue desarrollado a principios de los noventa, en Estados Unidos, para su uso en el Medical Outcomes Study (MOS, Estudio de los Resultados Médicos). Se trata de un cuestionario genérico que proporciona un perfil del estado de salud y es aplicable tanto a los pacientes como a la población general (169). Desde la primera publicación de su versión adaptada al castellano, se ha utilizado en numerosas investigaciones para la evaluación de resultados en salud (174). El SF-36 recoge dos dimensiones fundamentales de la CVRS, la física y la mental para los que ofrece dos puntuaciones resumen. Los pasos principales

INTRODUCCIÓN

de los algoritmos de cálculo de las puntuaciones de las 2 medidas sumario consisten en la estandarización de las 8 escalas del SF-36 (dolor corporal, función física, rol físico, salud general, vitalidad, función social, comportamiento emocional y salud mental) con la media y la desviación estándar (DE) de la población general y la ponderación de las escalas aplicando los pesos factoriales obtenidos en un análisis de componentes principales, y, finalmente, la agregación de las escalas y transformación para obtener una media de 50 y una DE de 10 en la población general (175).

Los estudios publicados sobre las características métricas de la versión española del SF-36 aportan suficiente evidencia sobre su fiabilidad, validez y sensibilidad. Estos resultados son consistentes con el instrumento original, lo que demuestra que la versión española es equivalente y puede ser utilizada en estudios nacionales e internacionales (174). Por otra parte, es importante destacar que se trata del cuestionario genérico para la medida de la CVRS que goza de mayor vitalidad. Desde su aparición en el inicio de la década de los noventa no ha cesado el diseño de nuevas herramientas de ayuda a la interpretación (valores de referencia poblacionales, valores sumario y diferencia).

Aunque el cuestionario SF-36 puede ser cumplimentado en un período de tiempo que oscila entre 5 y 10 min, contestar las 36 preguntas de respuesta múltiple, puede resultar demasiado tedioso para algunas personas y en determinadas situaciones. Para solventar este inconveniente surge el SF-12 (170) que es la forma reducida del SF-36 a 12 preguntas, que permite un cálculo global, pero no la medición de los diferentes aspectos de salud. Su uso estaría dirigido a evaluar conceptos generales de la salud física y mental en estudios en los que el uso del SF-36 fuera menos recomendable.

El SF-12 está formado por un subconjunto de 12 ítems del SF-36 obtenidos a partir de regresión múltiple, incluidos 1-2 ítems de cada una de las 8 escalas del SF-36. La información de estos 12 ítems se utiliza para construir las medidas sumario física y mental. Los ítems del cuestionario SF-12 explicaron más del 90% de la varianza de los índices sumario físico y mental del SF-36 en Estados Unidos y el 91% en España.

Aunque el SF-12 conlleva una pérdida de la precisión en las puntuaciones con respecto al SF-36, para estudios de comparación de grupos numerosos, estas diferencias no son tan importantes dado que los intervalos de confianza de las medias de los grupos están muy determinados por el tamaño de la muestra. Así pues, el SF-12 ha demostrado ser una alternativa útil al SF-36 cuando

se pretende medir la salud física y mental en general y el tamaño de muestra es elevado (500 individuos o más) (176).

Como versión aun más reducida del SF-36 nos encontramos el SF-8 , una versión de 8 preguntas que contiene un único ítem para cada una de las escalas del SF-36. Permite el cálculo de un perfil que fue construido para sustituir al SF-36 y el SF-12 en estudios de salud poblacionales en Estados Unidos y en el ámbito internacional (177).

El otro gran grupo de herramientas de medida de la CVRS lo constituyen los instrumentos específicos. Los instrumentos específicos se centran en aspectos específicos del estado de salud, e incluyen dimensiones de CVRS de una determinada patología, capacidad funcional, o de un determinado grupo de pacientes. Normalmente incluyen preguntas sobre el impacto de los síntomas. Se caracterizan por ser clínicamente más sensibles, pueden ser más discriminativos, pero no permiten comparación entre poblaciones y tienen un valor limitado en intervenciones poblacionales.

Son ejemplos de instrumentos específicos: ALQoL 9 para el alcoholismo (178) , AQLQ para el asma (179), EORTC QLQ-C30 para el cáncer (180), Dermatology Life Quality Index para enfermedades de la piel (181), ADDQoL para la diabetes (182), Pain Severity para el dolor (183), SGRQ y CCQ para pacientes con EPOC (184,185), MINICHAL para la hipertensión (186), el Minnesota LHFQ para la insuficiencia cardiaca (187), o el ECOS-16 para la osteoporosis (188).

Los instrumentos específicos para evaluar la CVRS en la artrosis de miembro inferior se revisarán en el siguiente epígrafe.

Un tercer tipo de instrumentos para medir la CVRS serían aquellos que miden de manera específica las preferencias de los sujetos por un determinado estado de salud.

A esa preferencia se le denomina “utilidad” y se le otorga un valor medio constante en una escala que va de 0 a 1; donde 0 se identifica con el peor estado de salud posible, la muerte, y 1 con la salud perfecta. En nuestro medio solo dos cuestionarios genéricos tienen la conversión directa en preferencias por los estados de salud que describen el SF-6D (175) y el EQ-5D, tanto en su versión 3L (189), como en la versión 5L (190). Algunos otros cuestionarios específicos de CVRS han desarrollado mecanismos de conversión de sus resultados en equivalencias con las utilidades

INTRODUCCIÓN

atribuidas a los estados descritos por el SF-6D o el EQ-5D , mediante un proceso denominado “mapping” (191).

El EQ-5D utilizaba una función de puntuación basada en preferencias para convertir el sistema descriptivo en una puntuación índice de resumen (189). Como se ha señalado se consideraba que el índice de utilidad debía oscilar entre 0, estado que equivalía a la muerte y 1 que representaba la salud perfecta. Pero en los estudios de preferencias de los diferentes estados de salud se han encontrado estados menos preferibles que la muerte para la población general, lo que hace que este índice pueda tomar valores negativos (hasta -0,421), al utilizar los algoritmos publicados para España (190). Dado que las utilidades son la herramienta necesaria para la construcción del AVAC, el EQ-5D es el cuestionario más ampliamente utilizado en el análisis de coste-efectividad de las tecnologías de atención de la salud (192).

Algunas organizaciones de evaluación de tecnologías de la salud europeas, como el National Institute for Health and Care Excellence (NICE) en el Reino Unido solicitan específicamente que el EQ-5D se utilice en todas las evaluaciones económicas que se les presenten para asegurar la comparabilidad entre los estudios (193,194). El EQ-5D en su última versión (EQ-5D-5L) es la alternativa recomendada para ser utilizada en las evaluaciones de tecnologías sanitarias en nuestro medio (190). El grupo autor del EQ-5D, Euroqol Group, ha propuesto que el EQ-5D tenga un papel en entornos clínicos, como parte de las ayudas para la decisión del paciente que pueda ser incluido entre los datos de rutina en la prestación de servicios de salud, e incorporado en las decisiones de atención médica (por ejemplo, las decisiones de los pacientes sobre qué tratamientos elegir) (195) .

El SF-6D, como decíamos anteriormente, es otro de los instrumentos de medida de la salud basado en preferencias. Obtenido a partir del SF-36, el SF-6D consiste en un sistema de clasificación del estado de salud de seis dimensiones, con cuatro a seis niveles en cada dimensión, lo que permite 18.000 estados de salud posibles (196). Las dimensiones del SF-6D incluyen: funcionamiento físico, limitaciones de rol, funcionamiento social, dolor, salud mental y vitalidad. El SF-6D se puede considerar como un medida de resultado anclada en una escala completa que abarca desde el valor 0 (identificado con la muerte) hasta 1 (salud completa) , donde los rangos teóricos plausibles comienzan desde 0,315 (el peor estado de salud posible) hasta 1 (salud completa), según el algoritmo de puntuación específico de la población (196).

2.8. Medición de la calidad de vida relacionada con la salud en pacientes con artrosis de miembro inferior

Las medidas de la CVRS son importantes no solo para evaluar la carga de la enfermedad o los resultados de cualquier intervención, sino también para ayudar a la toma de decisiones informadas en la asignación de recursos de salud a menudo limitados (197).

Como se señaló al hablar de la carga de la enfermedad, la artrosis de cadera y rodilla provoca un impacto muy significativo en cuanto a merma de calidad de vida. Este deterioro para el clínico se hace patente en la pérdida de funciones, pero podemos encontrar un deterioro que abarca las múltiples dimensiones que conforman la CVRS.

En este escenario se hace imprescindible el uso de herramientas validadas que nos permitan medir la percepción del estado de salud del sujeto y conocer el verdadero impacto de la artrosis de miembro inferior, incluyendo la perspectiva del paciente. Para ello disponemos de cuestionarios específicos para medir la CVRS. Entre todas las herramientas disponibles hay instrumentos para evaluar la CVRS de manera conjunta, o que sirven para valorar tanto la percepción de CVRS en personas con artrosis de cadera como de rodilla, y otras específicas de cada articulación.

Se describen a continuación las herramientas que sirven para valorar la CVRS en pacientes con artrosis de cadera o rodilla.

El cuestionario Western Ontario and McMaster University Osteoarthritis Index (WOMAC) (198), es uno de los más utilizados para la valoración de la CVRS en pacientes con artrosis en la extremidad inferior. Es un cuestionario auto-completado, dirigido específicamente para pacientes con artrosis de cadera o rodilla. Su escala multidimensional comprende 24 ítems agrupados en 3 dimensiones: dolor (5 ítems), rigidez (2 ítems) y funcionalidad física (17 ítems). Emplea un escala tipo Likert con 5 respuestas posibles para explicar la intensidad de cada ítem (ninguno, leve, moderado, severo o extremo), por lo que cada ítems recibe un puntaje de 0 a 4. Los puntajes se suman y estandarizan a partir de 0 a 100 (de mejor a peor), de modo que cuanto mayor sea el

INTRODUCCIÓN

puntaje, peor será estado de salud del paciente. Este cuestionario ha sido adaptado y validado para el entorno español; así como sus propiedades psicométricas, de modo que lo convierten en un cuestionario válido, fiable y lo suficientemente sensible para detectar cambios relevantes en pacientes con artrosis de cadera o rodilla (199). Aunque el cuestionario WOMAC puede cumplimentarse en apenas 5 minutos, la dimensión de funcionalidad física consta de 17 preguntas, siendo muy superior al resto de dimensiones. Este hecho puede causar una baja tasa de respuesta, sobre todo en estudios en los cuales haya una importante carga para el paciente en términos del número de preguntas a responder y si las mediciones se llevan a cabo en varios momentos.

De esta modo encontramos una versión WOMAC reducida (200), con una reducción de la dimensión de la funcionalidad física a 7 ítems, en lugar de los 17 ítems del cuestionario original. La versión española de la escala reducida WOMAC ha sido validada para los pacientes que van a ser intervenidos de prótesis total de cadera o de rodilla (201).

La escala de Lequesne (202), utilizada para valorar CVRS en pacientes con artrosis de cadera o rodilla, es un instrumento administrado por un entrevistador para pacientes con artrosis de cadera y rodilla. Consta de 3 dimensiones para un total de 10 ítems: dolor o malestar (5 ítems), distancia máxima caminada (1 ítems) y actividades de la vida diaria (4 ítems). A partir de la suma de los de los diferentes ítems se establece un índice de severidad de la artrosis de cadera o rodilla. El índice de severidad de la enfermedad de la cadera (ISH) se estableció, validó y evaluó como una nueva prueba de evaluación para el ensayo de nuevos fármacos, así como para el seguimiento a largo plazo de los pacientes y para ayudar con futuras indicaciones de cirugía. El índice de gravedad de la enfermedad de la rodilla (ISK) se validó y evaluó mediante los mismos métodos estadísticos. Su valor en ensayos con fármacos antiinflamatorios no esteroideos (AINE) o analgésicos es menor que el valor de la ISH. Sin embargo, su uso sigue estando justificado para tal fin y para el seguimiento a largo plazo de la artrosis de rodilla. Los índices tienen 2 ventajas: están estructurados por separado para la artrosis de cadera y rodilla; y el mismo instrumento sirve como medida de gravedad (escala de discapacidad) y como herramienta de medición de resultados en los ensayos (203). Este es un índice bastante simple que generalmente no causa malentendidos y se puede completar en solo 3 o 4 minutos. Las propiedades psicométricas de este índice se consideran adecuadas (204), y se ha traducido traducido y validado en diferentes

países como Turquía (205), Corea (206), Singapur (207), Suiza (208), Francia (209), y recientemente al idioma bengalí (210).

El cuestionario de evaluación de salud, Health Assessment Questionnaire (HAQ), introducido en 1980 por el “Stanford Arthritis Center”, se encuentra entre los primeros instrumentos para medir la CVRS en pacientes con artrosis de miembro inferior (211). Se trata de un autocuestionario, utilizado principalmente en países anglosajones, y se ha aplicado para la valoración de todo tipo de patología reumática, incluida la artrosis de cadera y rodilla. Existen numerosos estudios que han demostrado su fiabilidad, validez y sensibilidad al cambio (212). Conviene recordar que el HAQ completo es un cuestionario muy extenso que incluye 5 dimensiones: malestar, discapacidad, toxicidad por fármacos y coste económico. También denominadas como las “cinco D”: malestar (discomfort), discapacidad (disability), fármacos (drugs), coste (dollars) y muerte (death); pero a su vez, cada una de ellas se divide en subdimensiones y éstas en diferentes componentes (213).

Dada la popularidad de la escala de la discapacidad ha convertido la denominación HAQ en sinónimo de la misma. El cuestionario de discapacidad del HAQ consta de 20 ítems que evalúan el grado de dificultad (discapacidad física) autopercibida para realizar 20 actividades de la vida diaria agrupadas en 8 áreas. Existe una versión adaptada a nuestro país (214) donde se concluye que es una herramienta fiable y válida, pero que, para su uso en España, es recomendable contar con un entrevistador. También se han utilizado una versión reducida del HAQ, conocida como HAQ Modificado (Modified Health Assessment Questionnaire, MHAQ), con solo 8 ítems, uno de cada área (215). Sin embargo aporta menos información y lo que es más importante, plantea problemas psicométricos (216).

La escala funcional de las extremidades inferiores (LEFS) (217), fue desarrollado en un grupo de pacientes con diversos trastornos musculoesqueléticos para valorar la función de las extremidades inferiores. Se compone de 20 ítems relacionados en su mayoría con la actividades físicas con una demanda creciente. Cada ítem se puntúa en un escala tipo Likert de 0 (dificultad extrema) a 4 (ninguna dificultad). El cuestionario ha sido validado en múltiples idiomas, holandes, árabe, italiano, turco, brasileño, alemán, persa, y recientemente al chino (218). También encontramos una versión traducida y validada al idioma español (219).

El Índice de Impacto, es un cuestionario diseñado originalmente para evaluar el impacto de los síntomas urinarios de la hipertrofia benigna de próstata en la calidad de vida (“Benign Prostatic

INTRODUCCIÓN

Hyperplasia Impact Index”) (220). Consta de 4 ítems que explora 4 dimensiones de la CVRS, malestar físico, preocupación por la salud, molestias e impedimento para realizar actividades cotidianas. Las puntuaciones más altas en la escala nos dicen que el impacto del problema de salud en el paciente es mayor. Recientemente encontramos una validación del uso del índice de impacto para pacientes con artrosis de cadera y rodilla, presentando buenas propiedades psicométricas en cuanto a la fiabilidad y validez (221).

El cuestionario de Resultados de Cadera y Rodilla (Hip and Knee Outcomes Questionnaire), es una herramienta desarrollada por diez sociedades científicas americanas vinculadas a la patología del aparato locomotor. El cuestionario está planificado para evaluar el impacto en la calidad de vida de cualquier problema relacionado con el sistema musculoesquelético. Consta de un total de 16 preguntas, distribuidas en 7 ítems, de los cuales 3 tienen 4 subelementos. Emplea una escala tipo Likert con respuestas que osculan entre 5 o 7 opciones (222). Este cuestionario ha sido adaptado y validado en su versión española. Se comprueba que la versión en castellano aplicada al ámbito de España es fiable, válida y sensible a los cambios. Pero presenta un alto porcentaje de ítems sin responder en tres de sus preguntas; que se asocia a un posible fallo en las instrucciones, por lo que recomiendan la revisión del médico para su correcta cumplimentación (223).

El cuestionario OAKHQOL (Osteoarthritis Knee and Hip Quality of Life), se diseñó específicamente para evaluar la calidad de vida de los pacientes con artrosis de cadera o rodilla. Su escala multidimensional comprende 40 ítems distribuidos en 5 dimensiones: actividad física, salud mental, dolor, apoyo social y funcionamiento social. También incluye 3 ítems independientes sobre relaciones, actividad sexual y vida profesional. La escala de respuesta de los diferentes ítems van desde 0 (peor) hasta 10 (mejor). Se obtiene una puntuación media de cada dominio, que luego se estandariza en una escala de 0 (peor calidad de vida) a 100 (mejor calidad de vida). El instrumento original fue desarrollado en francés, y sus propiedades psicométricas se evaluaron en pacientes que cumplían los criterios de Altman para la artrosis de cadera o rodilla (224). Encontramos una versión traducida al español, cuyas propiedades psicométricas fueron validadas para medir la calidad de vida relacionada con la salud en los pacientes con artrosis de miembro inferior. Únicamente, indicar que en la validez convergente, en los dominios sociales, el OAKHQOL mostró una correlación débil con el SF-36 hecho que se explica porque ambos cuestionarios no miden los mismos aspectos dentro de la dimensión social (225).

Una vez analizados los principales instrumentos para medir la CVRS de la artrosis de miembro inferior, vamos a estudiar cuestionarios específicos para la cadera y la rodilla, respectivamente.

La escala de Harris (Harris Hip Score, HHS) (226), fue formulada inicialmente para valorar la patología traumática de la cadera y a los diferentes métodos de tratamiento. Siendo uno de los instrumentos más utilizados para evaluar los resultados obtenidos tras artroplastia de cadera, incluso antes de ser validadas sus propiedades psicométricas (227,228).

El cuestionario consta de 4 dimensiones: dolor, función, amplitud de movimiento y ausencia de deformidad, y utiliza un rango de puntuación desde 0 (peor capacidad funcional posible) y 100 (mejor capacidad funcional posible). La puntuación global se obtiene por la suma de las puntuaciones obtenidas en cada dominio, con un peso muy desigual de los mismos. El dolor (hasta 44 puntos) y la función (hasta 47 puntos, divididos en funciones de marcha, hasta 33 puntos, y actividades diarias, hasta 14 puntos) son las dimensiones o dominios que reciben un mayor peso. A la deformidad le corresponden 4 puntos y a la amplitud de movimiento 5 puntos. En la versión traducida al español se encontraron algunos problemas relativos al cumplimiento de los estándares de adaptación transcultural (229). Las propiedades psicométricas de la versión en español fueron validadas en pacientes que reciben rehabilitación tras artroplastia de cadera, resultando un instrumento válido y sensible a los cambios, pero con algunos matices. Las subescalas de movimiento y deformidad presentan problemas de validez y sensibilidad a los cambios, que se debe probablemente a que en origen estas subescalas se diseñaron para valorar pacientes provenientes de accidentes de tráfico con grandes deformidades, un problema infrecuente en la artroplastia, por lo que la casi totalidad de los pacientes obtienen la puntuación máxima (efecto techo) y muestran una baja validez de contenido. Actualmente estas dimensiones parecen innecesarias para la evaluación de pacientes tras artroplastia, pero dado el escaso valor en la puntuación global (4 y 5 puntos, sobre 100) no invalidan el cuestionario (229).

El cuestionario HOOS (Hip dysfunction and Osteoarthritis Outcome Score), se desarrolló como una extensión del WOMAC (198) , que se ha utilizado en todo el mundo como un cuestionario específico de artrosis de cadera. Consta de 40 ítems, distribuidos en 5 dimensiones: dolor (10 ítems), síntomas que incluyen rigidez y rango de movimiento (5 ítems), limitaciones de actividad diaria (17 ítems), función de deporte y recreación (4 ítems); y calidad de vida relacionada con la cadera (4 ítems). Las respuestas siguen una escala tipo Likert de 0 a 4 (no, leve, moderado, severo, extremo). El HOOS contiene todas las preguntas del cuestionario WOMAC, por lo que las

INTRODUCCIÓN

puntuaciones de WOMAC se pueden calcular a partir del cuestionario HOOS. La dimensión actividad de la vida diaria de HOOS es equivalente a la de funcionalidad física en el WOMAC. La versión original del HOOS demostró ser válida, fiable y sensible en pacientes con artrosis de cadera, y se considera útil para la evaluación de resultados relevantes para el paciente después de una artroplastia total de cadera (230). La mayoría de estos estudios de validación incluyen los dominios de deportes y recreación, ítems poco relevantes para la valoración de pacientes sometidos a cirugía de reemplazo articular. Estudios recientes han validado las subescalas de dolor y actividades de la vida diaria del cuestionario HOOS (y también del KOOS, cuestionario que más adelante describimos) en pacientes intervenidos de artroplastia, de modo que los resultados obtenidos se correlacionarán con los cambios en la calidad de vida informada por el paciente y su grado de satisfacción (231).

Encontramos una versión reducida a 12 preguntas, el HOOS – 12, que contiene los dominios: dolor (4 ítems), función (4 ítems), calidad de vida (4 ítems), y una puntuación resumida de impacto en la cadera (232). A pesar de ser uno de las herramientas más utilizadas para medir la CVRS en el paciente con artrosis de cadera no ha sido validado en la población española hasta el momento (233).

El Hip Outcome Score (HOS) (234), es un cuestionario autoinformado desarrollado en Pittsburgh (EE.UU) con el objetivo de evaluar los resultados de las intervenciones para el tratamiento de pacientes jóvenes con problemas de cadera. Consta de 28 ítems, divididos en dos escalas, actividades de la vida diaria (19 ítems) y la capacidad para realizar deporte, desde incapaz hasta sin dificultad (9 ítems). Ha sido validado para medir los resultados después de la artroscopia de cadera, destacando que la escala es específica para la evaluación de la función y no para la salud mental (235). La versión en español cumple las normativas de adaptación transcultural y presenta una evidencia fuerte de que el cuestionario HOS es una herramienta válida para la medición de resultados en el paciente joven con trastornos de cadera (236).

The International Hip Outcome Tool (iHOT-33) (237), es un instrumento autoadministrado para medir la CVRS en pacientes jóvenes y activos con trastornos de cadera. Consta de 33 ítems, distribuidos en 4 dimensiones: síntomas y limitaciones funcionales, deportes y actividades físicas recreativas, preocupaciones relacionadas con el trabajo, y preocupaciones sociales, emocionales y de estilo de vida. Las respuestas son valoradas utilizando una escala analógica visual. Las propiedades psicométricas de iHOT-33 muestran al cuestionario como una herramienta fiable,

válida y sensible al cambio para la evaluación del dolor de cadera en pacientes físicamente activos de 18 a 60 años. Hasta el momento en cuestionario ha sido validado al chino (238), holandés (239), alemán (240), francés (241), y también al español para pacientes sometidos a artroscopia de cadera (242). Encontramos una versión reducida del cuestionario, el iHOT-12 (243), que consta de 12 ítems que proceden del iHOT-33 original distribuidas en 4 dimensiones. La versión reducida ha sido validada en múltiples entornos culturales como el portugués, sueco, holandés, alemán, japonés y turco, y, recientemente, se ha validado la versión para Grecia del mismo (244), pero no así al español.

El Copenhagen Hip and Groin Outcome Score (HAGOS) (245), es un instrumento multidimensional desarrollado específicamente para su uso en pacientes jóvenes y de mediana edad con dolor crónico de cadera o ingle. Consta de 37 ítems, distribuidos en 6 dimensiones: síntomas (7 ítems), dolor (10 ítems), función física en la vida diaria (5 ítems), función física en el deporte y recreación (8 ítems), participación en actividades físicas (2 ítems) y calidad de vida relacionada con la cadera y / o la ingle (5 ítems). Encontramos versiones validadas al sueco (246), holandés (247) y chino (248), pero no encontramos versiones validadas al español en el momento actual.

El cuestionario Oxford Hip Score (OHS) (249), es un cuestionario autoadministrado que el paciente puede completar mediante una entrevista personal o enviarlo por correo una vez completado. Consta de 12 ítems con 5 respuestas posibles cada una, para evaluar la percepción del paciente durante las últimas 4 semanas. Se ha empleado con pacientes que padecen artrosis de cadera, tanto para estudiar su estado basal como para evaluar los cambios después de someterse a una cirugía de reemplazo de cadera. Cada respuesta recibe una puntuación de 0 a 4, donde 4 es el mejor resultado posible (250). Se obtiene una puntuación total de la suma de todas las respuestas, que van de 0 a 48, donde 48 es el mejor resultado posible. El OHS fue diseñado para evaluar el impacto de la cirugía de reemplazo total de cadera y se encontró que era más preciso que otros cuestionarios genéricos para ese propósito (251). Debido a sus buenas propiedades psicométricas, se ha comparado favorablemente con otras herramientas de uso generalizado que son más difíciles de administrar (252). Se ha adaptado al neerlandés (253), francés (254), alemán (255), italiano (256), danés (257), turco (258), coreano (259), japonés (260), rumano (261), y recientemente al árabe (262). El cuestionario también se adaptó al español en Colombia y se validó parcialmente, aunque no se verificó ni su sensibilidad a los cambios ni su estructura

INTRODUCCIÓN

factorial (263). De todas las adaptaciones mencionadas, la estructura factorial solo ha sido validada mediante un análisis factorial confirmatorio en la versión original del OHS.

El cuestionario de Dolor por Osteoartritis Intermitente y Constante, ICOAP (Intermittent and Constant Osteoarthritis Pain) ((264), fue desarrollado por un grupo de trabajo internacional, bajo la dirección de la Sociedad Internacional de Investigación sobre Osteoartritis (OARSI), para evaluar las dos formas de dolor reportadas por los pacientes con artrosis de cadera o rodilla: un dolor de fondo constante y uno menos frecuente, pero más intenso y a menudo impredecible. La herramienta unidimensional, esta compuesta por un total de 11 ítems, distribuidos en dos grupos, uno sobre dolor constante (5 ítems) y otro sobre dolor intermitente (6 ítems). Para cada uno de estos tipos de dolor, los ítems individuales evaluaron la intensidad del dolor, el efecto sobre el sueño, el impacto en la calidad de vida, el grado en que el dolor "frustra o molesta" y el grado en el que "preocupa o molesta". Para el dolor intermitente, se añadieron dos elementos adicionales para valorar la frecuencia del dolor y el grado en que se podía predecir el dolor. El período de tiempo utilizado fue de 1 semana. Cada respuesta recibe una puntuación de 0 nada a 4 extremo, para la intensidad del dolor; y de 0 nunca a 4 muy a menudo, para la frecuencia del mismo. El cuestionario ha sido validado a varios idiomas, el turco, portugués, alemán, griego y más recientemente el chino tradicional (265–269).

El formulario de Evaluación Subjetiva de la Rodilla del Comité Internacional de Documentación de la Rodilla (International Knee Documentation Committee, IKDC), se constituyó en 1987 por un grupo de cirujanos de rodilla de Europa y América. El cuestionario del IKDC fue inicialmente diseñado para lesiones ligamentosas de la rodilla (270), siendo sometido a varias revisiones, hasta que la Sociedad Ortopédica Americana de Medicina Deportiva realiza una nueva revisión con la publicación de un nuevo cuestionario específico de rodilla del IKDC (271), que fue validado para una amplia variedad de lesiones de rodilla, incluyendo lesiones ligamentosas, meniscales y del cartílago articular. El cuestionario consta de 18 ítems, con 3 dominios, incluidos síntomas, la función y la capacidad para participar en varios niveles de actividades deportivas para pacientes con una variedad de problemas de rodilla. El cuestionario ha sido validado en diferentes idiomas, italiano (272), griego (273), holandés (274), tailandés (275), brasileño (276), chino (277), coreano (278), persa (279), turco (280), y recientemente al japonés (281). Hasta el momento no disponemos de una versión validada al español.

El cuestionario KOOS (Knee Injury and Osteoarthritis Outcome Score), (282), se desarrolló como una extensión del cuestionario WOMAC para evaluar los síntomas y la función a corto y largo plazo de lesiones de rodilla y/o artrosis de rodilla. Consta de 42 ítems, distribuidos en 5 dimensiones: dolor (9 ítems), síntomas (7 ítems), actividades de la vida diaria (17 ítems), función deportiva y recreativa (5 ítems), y calidad de vida relacionada con la rodilla (4 ítems). Las respuestas se distribuyen según una escala tipo Likert de 0 a 4, y cada uno de los dominios se calcula como la suma de los ítems. La medida genera cinco puntuaciones por separado donde cuanto mayor es la puntuación, mejor es el estado de salud. El cuestionario ha sido validado al sueco (282), francés (283), japones (284), portugués (285), persa (286), holandés (287), inglés y chino de Singapur (288). Encontramos una versión traducida al español, en la que las propiedades psicométricas del KOOS español son similares a las del cuestionario original para su uso en pacientes con lesión condral sintomática de rodilla sometidos a intervención quirúrgica (289).

La versión reducida del cuestionario, el KOOS-Physical Function Short form (KOOS-PS) surge para valorar exclusivamente la función física en la artrosis de rodilla (290). Consta de solo 7 ítems derivados de las actividades de la vida diaria y las subescalas de función deportiva y recreativa.

El cuestionario KSS (Knee Society Score), el sistema de clasificación clínica de la Sociedad Americana de Rodilla, se desarrolló en 1989 para evaluar la capacidad funcional de los pacientes antes y después de ser sometidos a cirugía de reemplazo articular (291). Con el tiempo, se ha cuestionado la fiabilidad, capacidad de respuesta y la validez de la herramienta original en los pacientes actuales sometidos a cirugía protésica, los cuales tienen unas expectativas, demandas y requisitos funcionales muy diferentes a los pacientes que inicialmente se sometían a cirugía de artroplastia de rodilla (292,293).

En 2012, se introdujo un nuevo "Knee Society Knee Scoring System" (KSS) para satisfacer la necesidad de un sistema de puntuación que caracterice mejor las expectativas, la satisfacción y las actividades físicas de una población actual (292).

En cuestionario consta de dos versiones, una para administrarse antes de la intervención y otra después de la cirugía. La puntuación combina un componente objetivo derivado del médico y un componente informado por el paciente. La puntuación completada por el cirujano incluye 3 dominios principales: dolor, estabilidad y rango de movimiento, a los que se añaden otros 3 dominios: contractura en flexión, pérdida de extensión y alineación consideradas como deducciones a las puntuaciones anteriores. La puntuación de la función se ha mejorado con el

INTRODUCCIÓN

componente informado por el paciente, que incluye las actividades de la vida diaria, la práctica deportiva, actividades recreativas, la satisfacción y las expectativas. El análisis de la versión traducida al español y adaptación cultural del KSS demuestra que es válido, fiable y sensible al cambio para pacientes sometidos a artroplastia primaria de rodilla (294). Encontramos una versión reducida del KSS a 6 ítems que es recomendada para el uso general del médico en la práctica clínica, y que ha sido válida para evaluar el resultado funcional de los pacientes intervenidos de artroplastia de rodilla (295).

El cuestionario Oxford Knee Score (OKS), es un cuestionario breve autoadministrado desarrollado para medir el impacto de la cirugía de artroplastia total de rodilla en la percepción de CVRS por parte de los pacientes (296). Puede ser respondido mediante entrevistas “cara a cara” o enviadas por correo por el paciente una vez completado. Consta de 12 ítems, con 5 respuestas posibles cada una, destinadas a evaluar la percepción de calidad de vida del paciente durante las últimas 4 semanas. Se ha utilizado tanto para evaluar la situación basal en pacientes con artrosis de rodilla, como para estudiar los cambios después de ser sometidos a cirugía de reemplazo articular. Cada respuesta recibe una puntuación de 0 a 4, donde 4 es el mejor resultado posible (250) . La puntuación total es la suma de la obtenida en todos los ítems, que varía de 0 a 48, donde 48 es el mejor resultado posible. Se encuentra entre los cuestionarios con mayor validez, sensibilidad al cambio y fiabilidad para la valoración de personas con afección de rodilla (297). Debido a sus buenas propiedades psicométricas, se ha comparado favorablemente con otras herramientas ampliamente utilizadas en diferentes idiomas que son más difíciles de administrar (298,299). Hasta el momento el cuestionario OKS ha sido adaptado y validado en italiano (300), holandés (301), chino e inglés de Singapur (302), alemán (303), francés (304), japonés (305), portugués (306), coreano (307), persa (308), griego (309), español de Colombia (263), árabe (310) y finlandés (311).

2.9. Validación de herramientas para medir la calidad de vida relacionada con la salud

Como se ha señalado previamente, se reconoce de forma generalizada que el impacto que las enfermedades tienen sobre los pacientes no puede ser descrito en su totalidad por medidas de salud “objetivas” como las tradicionalmente consideradas, relativas a la morbilidad, la mortalidad o incluso a la demanda de uso de servicios sanitarios. La incorporación de la perspectiva del paciente a la valoración de su propio proceso de salud y enfermedad puso la CVRS como un resultado en salud que debe ser considerado y evaluado. Pero también se ha presentado la CVRS como un constructo complejo (148), que debe incorporar en su valoración la subjetividad del sujeto y que necesita de herramientas específicas para ser aprehendido.

La evaluación estandarizada de los aspectos más subjetivos relacionados con la salud aporta ventajas diversas a la comprensión de dicha situación de salud y también a la evaluación de los resultados sobre las intervenciones que sobre ella se realizan, pero existen inconvenientes ligados a la medición de conceptos intangibles, latentes y no siempre fáciles de observar (312).

Para poder aprehender correctamente esos conceptos intangibles se han desarrollado instrumentos de medida, conocidos genéricamente como cuestionarios. Un cuestionario se define como un instrumento para la medición de uno o más constructos por medio de puntuaciones agregadas de los ítems, denominadas escalas. Los ítems de un cuestionario suelen estar completamente estructurados, tienen un formato similar, y requieren un juicio o descripción por parte de un encuestador o evaluador (313). El proceso de comprensión e interpretación de los juicios subjetivos realizados por los individuos sobre su estado de salud, y su posterior incorporación a la toma de decisiones, precisa de un conocimiento detallado del uso de dichos instrumentos, que pasa necesariamente por la comprensión del proceso metodológico subyacente a su construcción y funcionamiento (312). La utilización generalizada de un cuestionario viene precedida de la comprobación de que la herramienta mide apropiadamente y de modo constante en las mismas circunstancias, que es sensible a los cambios de la situación clínica, que en la práctica no presenta dificultades de aplicación, etc. Toda esta comprobación de sus características es lo que se denomina proceso de validación (314).

INTRODUCCIÓN

La construcción de un cuestionario es un proceso extraordinariamente largo y minucioso que precisa de equipos multidisciplinares y de grandes recursos (313), por lo que, por motivos de eficiencia o de necesidad, como ocurre en estudios internacionales, se precisa utilizar herramientas creadas en otros idiomas o en otros entornos culturales. Para ello es necesario adaptar esos cuestionarios. La adaptación transcultural de un cuestionario sobre el estado de salud para su uso en un nuevo país, cultura y/o idioma requiere el uso de un método que asegure la equivalencia entre las versiones de origen y destino del cuestionario. El proceso de adaptación transcultural implica el desarrollo de versiones de un instrumento de evaluación que son equivalentes al original, pero al mismo tiempo, lingüística y culturalmente adaptadas a un contexto diferente al original. Por lo tanto, la adaptación de los instrumentos de evaluación permite comparar los resultados de las investigaciones realizadas en diferentes culturas, ayudando al intercambio de información dentro de la comunidad científica. Pero la adaptación de cuestionario debe terminar con un nuevo proceso de validación que asegure que las propiedades de medida de la versión original se mantienen en la adaptada (315).

El proceso de validación es imprescindible como se ha señalado para delimitar la adecuación del uso de un cuestionario, tanto si se ha creado "ad hoc", como si se ha adaptado de otro idioma o entorno cultural. El proceso de validación implica la aceptación del hecho de que se pueden utilizar medidas subjetivas para captar realidades o constructos complejos, en el caso que nos ocupa, la CVRS. Los argumentos para considerar los juicios subjetivos como un enfoque válido de medición provienen del campo de la psicofísica, que investiga las características del ser humano como instrumento de medición. Los principios psicofísicos se incorporaron posteriormente a la psicometría. La psicometría aplica métodos psicofísicos para medir cualidades para las que no existe una escala física, y esto constituye un elemento fundamental en el desarrollo de los métodos de medición de la salud (316).

El proceso psicométrico trata de asignar valores a las percepciones que expresan los sujetos (312). Los posibles sucesos observables que definen el constructo son innumerables, por lo que las cuestiones que componen la herramienta de medida, el cuestionario, pueden considerarse una muestra representativa del universo de preguntas que podían definir el constructo de interés y las respuestas de los sujetos a dichas preguntas o afirmaciones que éstos formulan presentan la concreción observable del constructo que está siendo evaluado (156,312). Existen varias formas de evaluar el desempeño de las herramientas de medida de las percepciones subjetivas, que se

engloban en dos grandes paradigmas psicométricos: la teoría clásica de los test (TCT, o Classic Test Theory, CTT) (317), y la teoría de respuesta al ítem (TRI, o Item Response Theory, ITR) (318).

La TCT se construye sobre el principio de que una puntuación observada está determinada por el estado real de la variable inobservable de interés, más el error aportado por todas las demás influencias sobre la variable observable. El estado real de la variable inobservable es su puntuación verdadera hipotética. Aunque no podemos acceder directamente a esa experiencia subjetiva, es razonable suponer, a partir de la lógica y de nuestras propias experiencias pasadas, que la experiencia puede indentificarse con cierto grado o cantidad. El investigador tiene que llegar a conclusiones sobre lo experimentado basándose en lo que el sujeto expresa. Hay muchas posibilidades de error en este proceso. La puntuación observada, la información que obtenemos del paciente en respuesta a una pregunta, tiene dos componentes, el verdadero nivel de esa variable y un error asociado a ese ítem concreto, es decir, la cantidad de influencia que ejercen sobre ese ítem todas las circunstancias que rodean a la expresión de la experiencia real. La variable que se observa es, por tanto, una mezcla de información relevante y error (319).

Los modelos de la TRI, a veces también llamado análisis de rasgos latentes, difieren de aquellos generados bajo la TCT en muchos aspectos. La primera diferencia es que la TRI no proporciona modelos de puntuaciones de test. En la TRI el objeto de interés es el parámetro del sujeto o la variable latente tal como se especifica en el modelo (318). El rasgo o variable latente se refiere a un continuo teórico que los ítems del test están diseñados para medir. Si estuviésemos midiendo la capacidad funcional, el rasgo latente sería el continuo que va de la capacidad plena a la discapacidad grave y se traza a lo largo del eje horizontal de un gráfico. Los ítems que miden diferentes grados de discapacidad estarían espaciados a lo largo de este continuo. Ítems como “puedo subir una cuesta pronunciada sin esfuerzo” estarían cerca de la izquierda de ese eje horizontal imaginario, y otros como “me fatigo al levantarme de la cama” estarían cerca del otro extremo de dicho eje (316). La tarea estadística consiste en estimar los parámetros del modelo y evaluar el ajuste del modelo a las respuestas de los ítems. La segunda diferencia está asociada a la primera. En la TRI la modelización no se dirige al conjunto del test como en la TCT, sino que se dirige al nivel del ítem. Se reconoce que los datos recogidos en las pruebas psicológicas son siempre respuestas cualitativas a los ítems del test. La información cuantitativa se obtiene a partir de los datos cualitativos mediante el uso de modelos de medición formulados con este fin. La TRI incorpora en el modelo la dificultad para responder a un ítem, esto es la posición del sujeto que

INTRODUCCIÓN

contesta. La TRI no ofrece una puntuación cuantitativa del test, sino que considera la información incluida en las respuestas de los ítems como meramente cualitativa y utiliza parámetros cuantitativos para explicarla. La escala de los parámetros viene dada por la estructura del modelo elegido. La tercera diferencia es que los modelos de respuesta al ítem son estocásticos. Las respuestas a los ítems se consideran resultados de procesos estocásticos que pueden caracterizarse por ciertos parámetros pero en los que, las perturbaciones aleatorias desempeñan su papel. En consecuencia, los modelos de respuesta al ítem no son modelos para la explicación de las respuestas al ítem, sino de las probabilidades de estas respuestas (318). La TRI está basada en los métodos de Rasch, un enfoque de escalamiento alternativo, propuesto por el matemático danés Georg Rasch (320) que asumen los siguientes supuestos: a) el atributo que se desea medir puede representarse en una única dimensión en la que se situarían conjuntamente las personas y los ítems; b) el nivel de la persona en el atributo y la dificultad del ítem determinan la probabilidad de que la respuesta sea correcta.

La elección del enfoque psicométrico que se elige depende de diferentes factores entre los que parece fundamental la facilidad de manejo para el investigador y la disponibilidad de software (321). En este trabajo utilizaremos como aproximación psicométrica general la TCT, salvo para realizar análisis factorial confirmatorio, por lo que haremos una descripción del proceso de validación bajo dicho paradigma.

Para evaluar las propiedades psicométricas de una herramienta de medida, de un cuestionario, nos fijaremos en los siguientes aspectos: la factibilidad, la validez, la fiabilidad, la sensibilidad al cambio y la interpretabilidad.

La factibilidad mide si el cuestionario es asequible para utilizarlo en el entorno que se quiere utilizar y se valora a través del tiempo que se tarda en cumplimentar, de las preguntas que se dejan sin responder y de la percepción de sujetos y profesionales (314).

La validez se define como el grado en que un instrumento de resultados informado por el paciente y relacionado con la salud (instrumento HR-PRO) mide el constructo o constructos que pretende medir. Tiene varios componentes (322):

- Validez de contenido: examina el grado en que los contenidos de interés están ampliamente representados por los ítems del cuestionario. Este concepto incorpora la

“validez aparente” o el grado en que los ítems del instrumento parecen ser un reflejo adecuado del constructo que se pretende medir. Para poder valorar la validez de contenido de un cuestionario los autores deben reportar, el objetivo de su uso, el concepto o constructo que quiere aprehender, la población a la que se aplica, los ítems que se utilizan y los que se han rechazado (323). Algunos autores consideran parte de la valoración de la validez de contenido el análisis de la existencia de los denominados “efecto suelo” y “efecto techo”. Estos se corresponden con el agrupamiento de la respuesta en la parte inferior (o superior) de la escala indicando un problema en el rango de respuesta ofrecido. La presencia de estos se relacionan con un aumento en la incertidumbre de los resultados, especialmente por el sesgo que producen en la capacidad discriminativa de la herramienta para encontrar diferencias entre grupos (324).

- Validez de constructo: mide el grado en que las puntuaciones de un instrumento concreto se relacionan con otras medidas de forma coherente con las hipótesis teóricas sobre los conceptos que se miden. Se evalúa poniendo a prueba hipótesis predefinidas (por ejemplo, sobre las correlaciones esperadas entre las mediciones o las diferencias esperadas en las puntuaciones entre grupos "conocidos"). Estas hipótesis deben ser tan específicas como sea posible (323).
- Validez de criterio: se evalúa por el grado en que las puntuaciones de un instrumento son un reflejo adecuado de un “patrón oro”.

La fiabilidad (“reliability” en la literatura internacional) se define como la medida en que las puntuaciones de los pacientes que no han sufrido cambios permanecen iguales tras mediciones repetidas en varias condiciones. La fiabilidad hace referencia a la ausencia de error de medida. Existen varios componentes de la fiabilidad (322):

- La consistencia interna que evalúa el grado de interrelación entre los ítems. Esta propiedad se evalúa a través de un análisis de componentes principales o un análisis factorial exploratorio, seguido de un análisis factorial confirmatorio. Cuando la consistencia interna es relevante, debe aplicarse el análisis de componentes principales o el análisis factorial para determinar si los ítems forman una sola escala global (dimensión) o más de una (323). Después de determinar el número de (sub)escalas, se debe calcular el alfa de Cronbach para cada una de ellas por separado.

INTRODUCCIÓN

- La reproducibilidad, es la proporción de la varianza total en las mediciones que se debe a las diferencias verdaderas entre los pacientes. Se puede evaluar la reproducibilidad de los resultados de un instrumento para catalogar a sujetos iguales a lo largo del tiempo (fiabilidad test-retest), resultados obtenidos por parte de diferentes personas en la misma ocasión (fiabilidad interobservador) o por las mismas personas en diferentes ocasiones (fiabilidad intraobservador). La reproducibilidad se refiere al grado en que los pacientes pueden distinguirse entre sí, a pesar del error de medición. Esta característica es importante para fines discriminatorios si se quiere distinguir entre pacientes, por ejemplo, con una enfermedad más o menos grave. Se evalúa a través del coeficientes de correlación intraclass que estudia la variación en la población (variación interindividual) dividida por la variación total, que es la variación interindividual más la variación intraindividual (error de medición), expresada como una relación entre 0 y 1.

- El error de la medida, se refiere al error de medición absoluto, es decir, a la proximidad de las puntuaciones en las medidas repetidas, expresada en la unidad de la escala de medición en cuestión. Se explora a través del error estándar de medición (EEM, Standard Error of Measurement, SEM), que representa la cantidad de error asociada a la evaluación de un sujeto individual. Se requiere un pequeño error de medición para fines de evaluación en los que se desea distinguir los cambios clínicamente importantes del error de medición.

Tiene sentido recordar en este punto que la TCT asume que la puntuación de cada ítem es un reflejo de la verdadera puntuación del fenómeno de interés y del error. Además, el error se distribuye aleatoriamente en torno a una media de cero, lo que significa que los errores asociados a diferentes ítems tenderán a anularse entre sí. Una consecuencia de esa anulación es que el error tendrá poco efecto en las puntuaciones medias, pero aumentará las varianzas de los ítems. A medida que se incluyen más ítems en una escala, es más probable que los errores asociados a cada ítem individual se equilibren entre sí y, por tanto, tengan un efecto menor en la suma o puntuación media de los ítems. Esto es directamente análogo a la mejora que se consigue al estimar el valor medio de una población en alguna variable cuando se toma una muestra de un mayor número de individuos. En consecuencia, la TCT hace hincapié en que las medidas adquieren fuerza a través del número de ítems que comprenden. Un mayor número de ítems proporciona una mayor fiabilidad de la escala (319).

La sensibilidad al cambio, es una propiedad psicométrica de la herramienta de medida que se define como su utilidad para detectar cambios en el tiempo en el constructo evaluado (322). Los instrumentos para aprehender estados de salud, han sido clasificados en discriminativos, cuando su objetivo es encontrar diferencias entre sujetos, y evaluativos, si su pretensión es medir el cambio en el tiempo en el mismo individuo o en el mismo grupo (147). Un instrumento de carácter evaluativo, no estará correctamente validado hasta que no se haya determinado su sensibilidad al cambio. El instrumento debe ser capaz de distinguir los cambios clínicamente importantes de los errores de medición. Esta característica se puede valorar mediante la estimación de la mínima diferencia clínicamente importante (MDCI) y el mínimo cambio detectable (MCD). Estas dos medidas están relacionadas con la capacidad de respuesta, pero están más orientadas a la clínica y centradas en el nivel individual. La MDCI se estima con el cambio medio en los pacientes que han mejorado su estado en el tiempo. El MCD expresa la magnitud mínima de cambio por encima de la cual es probable que el cambio observado sea real y no sólo un error de medición. Para la estimación del MCD se utiliza el error estándar de medición (EEM). Cuando la relación entre la MDCI y la MCD es superior a 1 se puede decir que la MDCI detectada es relevante pues puede ser distinguida del error de medición (325).

La interpretabilidad, es el último elemento a valorar en el proceso de validación de un cuestionario y se define como el grado en que se puede asignar un significado cualitativo, es decir, connotaciones clínicas o comunes, a las puntuaciones cuantitativas de un instrumento o a los cambios en las puntuaciones (322).

La evaluación de todas estas propiedades de un instrumento que valoran un constructo complejo permiten que éste sea utilizado en distintas poblaciones y circunstancias y que sus resultados sean comparables. En el caso de las herramientas que valoran la CVRS, los procesos de validación son imprescindibles para poder asegurar que estamos evaluando el constructo adecuado en la forma adecuada. La validación de un cuestionario, para medir la CVRS o para otros propósitos, es un proceso complejo que implica múltiples fuentes de información y la recogida de diferentes evidencias empíricas. Dado que no existe un criterio concreto ni único a partir del cual pueda considerarse que un cuestionario sea válido, podemos entender la validación de una herramienta como un proceso dinámico (326), pues los instrumentos van adquiriendo consistencia en su

INTRODUCCIÓN

validez con su uso y no se debe renunciar a medir sus propiedades psicométricas cuando se utilizan en circunstancias que puedan aportar diferencias valorativas (314).

3. Justificación

JUSTIFICACIÓN

Como se ha discutido previamente, la artrosis es una enfermedad de origen multifactorial y curso variable (2), extraordinariamente presente en la población general. Su prevalencia poblacional puede llegar a más del 10 % (100), aumentando con la edad y siendo diagnosticada con más frecuencia en países desarrollados (327).

Cuando nos referimos a la artrosis del miembro inferior (OAMI), su prevalencia, también es muy significativa, del 3.8%, para la rodilla y del 0,85% para la cadera a nivel mundial (119). En España, además de una elevada prevalencia de artrosis de miembro inferior, se ha reportado una notable incidencia, comunicándose tasas de incidencia de 6,5 y 2,1 /1000 personas-año para la artrosis de rodilla y de cadera definidas clínicamente (328).

Aunque al estandarizar la prevalencia por edad, algunos autores no encontraron incrementos en la prevalencia de artrosis de miembro inferior entre 1990 y 2010 (119), estudios posteriores han evidenciado incrementos en las tasas de prevalencia ajustadas por edad del 1,8% y el 2,3 % para la artrosis de cadera y rodilla, respectivamente entre el 2005 y 2015 (121). Para el caso particular de la artrosis de rodilla, en algunas cohortes se ha observado un aumento de la prevalencia de la gonartrosis sintomática de alrededor 10% en los varones y del 6% en las mujeres de la cohorte Framingham (122). Otros estudios, que amplían el periodo de observación, reportan que la prevalencia de artrosis de rodilla se ha duplicado en Estados Unidos desde mediados del siglo pasado sin que este incremento pueda explicarse solo por una mayor longevidad de la población (329).

La artrosis de miembro inferior es la mayor causa de dolor y discapacidad funcional en la población mundial (330) y su impacto en forma de carga de enfermedad es extraordinario, suponiendo una pérdida en años de vida ajustados por discapacidad (AVAD) del 0,69% de todos los AVAD perdidos en la población mundial en 2010 (119). Este impacto de la artrosis de miembro inferior en forma de carga de enfermedad ha crecido, de manera que entre 2005 y 2015 los AVAD perdidos se incrementaron un 34,6% para la artrosis de rodilla y un 35,8% para la artrosis de cadera (121).

Además del impacto social que tiene la enfermedad, desde la perspectiva individual, la artrosis de miembro inferior tiene una enorme repercusión en la vida del paciente. Esta repercusión puede intentar valorarse con herramientas clásicas, como las que miden la morbilidad relativas al dolor y/o a la dificultad funcional, pero se sabe que esta condición clínica tiene implicaciones sobre otros campos de la salud, como aquellos relativos a la salud mental (131,132). Por ello cobra especial relevancia el valorar el impacto de la artrosis de miembro inferior sobre la vida del paciente mediante herramientas que incorporen la perspectiva del propio paciente.

JUSTIFICACIÓN

La evolución de los conceptos de salud y enfermedad han hecho necesaria la incorporación de medidas de la salud que incorporen todas las dimensiones que incorpora el constructo de salud en nuestros días. De hecho cobran importancia las medidas subjetivas de salud, pues aparecen como las más sensibles para la detección de deterioros que son primeramente percibidos por el propio paciente. Los avances metodológicos en la medición de la percepción de la salud han permitido incorporar esa perspectiva del paciente, mediante herramientas que valoran el estado de salud, o la CVRS (146). La aprehensión de la percepción que tiene el paciente de su salud mediante la evaluación de la CVRS permite caracterizar la enfermedad, especialmente aquella con características de cronicidad, como ocurre con la artrosis de miembro inferior, y esta caracterización tiene implicaciones para la comprensión del sujeto individual, para el abordaje de su situación concreta, pero también para la planificación de la salud poblacional, para la evaluación de las intervenciones en salud y para la planificación en salud desde la meso o la macrogestión (147,158). Además se conoce que la percepción del estado de salud está impregnado de elementos culturales y sociales (149), que hacen necesario que la medición del impacto de determinada condición clínica sobre la salud se haga en el contexto en el que luego se van a tomar las decisiones sobre el abordaje de esa enfermedad o condición .

Como se expuso en la Introducción, las herramientas para evaluar CVRS pueden ser genéricas o específicas de enfermedades concretas. Las herramientas genéricas tiene un alto grado de desarrollo y utilización, pero aquellos instrumentos específicos de determinada enfermedad o condición clínica pueden ser más útiles para discriminar la situación de sujetos con dicha enfermedad y más sensibles a los cambios que se produzcan (146). De ahí el interés de contar con instrumentos específicos para evaluar la percepción de la CVRS específicos de enfermedades prevalentes y que impactan en múltiples dimensiones de la salud, como ocurre en el caso de la artrosis de miembro inferior.

Pero el uso de estas herramientas requiere de un proceso complejo que implica la generación y análisis de diversas evidencias empíricas, que aseguren que son útiles para el propósito para el que van a utilizarse. Este proceso no se sustenta en un criterio concreto para asegurar la idoneidad de una herramienta sino que está dotado de un dinamismo que aporta consistencia en su validez, según se avanza en su utilización, por lo que debe perseguirse la evaluación de las propiedades psicométricas de la citada herramienta siempre que se pueda aportar alguna diferencia valorativa (314,326).

Caracterización de la Calidad de Vida Relacionada con la Salud en pacientes con artrosis de miembro inferior y validación de nuevas herramientas para su evaluación.

Estos condicionantes nos llevan a encontrar pertinente el estudio del impacto sobre la salud, desde la perspectiva del paciente, de la artrosis de miembro inferior y la evaluación de nuevas herramientas de valoración de la CVRS, específicas para esta condición clínica.

JUSTIFICACIÓN

4. Hipótesis y objetivos

HIPÓTESIS Y OBJETIVOS

Hipótesis

El impacto de la artrosis de miembro inferior sobre la percepción de Calidad de Vida Relacionada con la Salud (CVRS) expresada por los sujetos que la sufren es muy significativa, y puede ser detectada mediante la utilización de instrumentos generales.

Existe un nuevo instrumento de medida de la CVRS en pacientes con artrosis de cadera, la versión adaptada al castellano del Oxford Hip Score, cuyas propiedades psicométricas le hacen útil para detectar diferencias entre sujetos y para valorar la evolución de personas que sufren la enfermedad o de grupos de sujetos con dicha condición clínica.

Existe un nuevo instrumento de medida de la CVRS en pacientes con artrosis de rodilla, la versión adaptada al castellano del Oxford Knee Score, cuyas propiedades psicométricas le hacen útil para detectar diferencias entre sujetos y para valorar la evolución de personas que sufren la enfermedad o de grupos de sujetos con dicha condición clínica.

HIPÓTESIS Y OBJETIVOS

Objetivos

Objetivo general

Estimar el impacto de la artrosis de miembro inferior en la percepción de la Calidad de Vida Relacionada con la Salud (CVRS) en la población general y valorar las propiedades psicométricas de los instrumentos de medida específicos Oxford Hip Score y Oxford Knee Score, en sus versiones adaptadas al castellano.

Objetivos específicos

1. Evaluar el impacto de la artrosis de miembro inferior en la Calidad de Vida Relacionada con la Salud percibida por los sujetos que la sufren, comparándola con la situación de la población general, ajustando por características sociodemográficas y otras condiciones clínicas.
2. Comparar las diferencias en la percepción de la CVRS entre sujetos que sufren artrosis de cadera o rodilla.
3. Estudiar potenciales características individuales y contextuales que se asocien con la percepción de la CVRS en pacientes con artrosis de miembro inferior.
4. Evaluar las características de fiabilidad y validez del cuestionario Oxford Hip Score, adaptado al castellano en nuestro medio.
5. Valorar la sensibilidad al cambio del cuestionario Oxford Hip Score, adaptado al castellano en nuestro medio.
6. Evaluar las características de fiabilidad y validez del cuestionario Oxford Knee Score, adaptado al castellano en nuestro medio.
7. Valorar la sensibilidad al cambio del cuestionario Oxford Knee Score, adaptado al castellano en nuestro medio.

HIPÓTESIS Y OBJETIVOS

5. Material y Métodos

5.1. Diseño del estudio.

Se planteó un estudio observacional longitudinal multicéntrico, con enfoque analítico, con un seguimiento de seis meses.

Para dar respuesta a los objetivos específicos del 4 al 7 se realizó un diseño de validación de un cuestionario.

5.2. Ámbito

Se presentan el ámbito temporal y el geográfico en el que se llevó a cabo el estudio.

Ámbito temporal.

La inclusión de sujetos en el estudio se hizo entre el 1 de enero y el 31 de diciembre de 2015.

El seguimiento de los sujetos se prolongó hasta el 30 de junio de 2016.

Ámbito geográfico.

Los sujetos de estudio fueron reclutados en tres consultas de tres centros de salud y dos consultas de reumatología y dos de traumatología de dos hospitales de Canarias; en doce centros de salud y dos consultas de traumatología de dos hospitales en Vizcaya; y en ocho consultas de seis centros de salud y cuatro consultas de traumatología en dos hospitales en Madrid.

En total participaron en el reclutamiento veintidós centros de atención primaria y seis centros hospitalarios de los tres territorios mencionados.

5.3. Población de estudio y criterios de selección

La población de estudio pretendía ser representativa de aquellos sujetos con artrosis de cadera o rodilla que son diagnosticados en cualquier ámbito en el Sistema Nacional de Salud.

Para delimitar dicha población de estudio se utilizaron los procedimientos que se describen a continuación.

Fueron criterios de inclusión en el estudio:

- Pacientes mayores de 18 años.
- Pacientes atendidos en los servicios mencionados de los centros participantes en este proyecto.
- Pacientes diagnosticados de artrosis de cadera y/o rodilla.
- Pacientes que tras ser informados del objetivo del estudio aceptaron participar.

El diagnóstico de artrosis de cadera se hizo mediante los criterios de la American College of Rheumatology (ACR), que ofrecen una sensibilidad del 89% y una especificidad del 91% (331). Estos criterios se refieren a la existencia de dolor cadera y al menos 2 de las siguientes circunstancias: VSG < 20 mm/h, osteofitos en la radiografía o estrechamiento del espacio articular en la radiología.

Para el diagnóstico de artrosis de rodilla también se utilizaron los criterios diagnósticos de la American College of Rheumatology (ACR), que ofrecen una sensibilidad del 91% y una especificidad del 86% (115). Dichos criterios incluyen reportar dolor en rodilla y presentar osteofitos en un estudio radiológico y, al menos uno de los tres siguiente: edad mayor de 50 años, rigidez matutina menor de 30 minutos y crepitación.

Los criterios de exclusión utilizados en el estudio fueron los siguientes:

- Tener patologías orgánicas o psiquiátricas que impidiesen la colaboración en el estudio.
- Pacientes analfabetos o con déficit sensoriales que no permitiesen completar los instrumentos de medición de resultados incluidos en el estudio.

5.4. Tamaño muestral y selección de la muestra.

Los sujetos fueron incluidos mediante un muestreo de conveniencia entre aquellos que cumplían los criterios de inclusión, no presentaban criterios de exclusión y daban su consentimiento escrito a participar.

Para explicar la estimación del tamaño muestral se deben tener en cuenta los diferentes objetivos específicos. Como se explicará más adelante la respuesta a los tres primeros objetivos específicos se realizará mediante la construcción de modelos explicativos. El tamaño muestral necesario para un modelo que ofrezca estimaciones fiables depende del número de variables explicativas. Se requieren al menos quince observaciones por variable explicativa, considerando cada una de las posibles variables “dummy”(332). Dado que no se incluirán más de 20 variables en ningún modelo, incluidas las variables “dummy” un tamaño muestral de 300 sujetos permitía alcanzar estos objetivos.

En referencia a los objetivos relacionados con las características psicométricas de los instrumentos para medir CVRS, la técnica más exigente en cuanto a tamaño muestral era el análisis factorial confirmatorio. Se calculó que con 300 pacientes se podía realizar un análisis factorial confirmatorio para un cuestionario con un solo factor de 12 ítems (333). Con ese tamaño además se podían estimar coeficientes de correlación intraclase (CCI) superiores a 0,8 con una precisión inferior al 10% (334).

Pero este sería el tamaño requerido para cada uno de los grupos a estudio, sujetos con artrosis de cadera y sujetos con artrosis de rodilla. El tamaño muestral final debía ser de aproximadamente 600 sujetos que completasen el seguimiento de 6 meses. Asumiendo unas pérdidas durante el seguimiento inferiores al 10%, se necesitaba reclutar alrededor de 670 pacientes.

Para dar respuesta al objetivo específico 1 se necesitaba un grupo de comparación representativo de la población general. Se escogió una muestra representativa de la población española de 50

años, cuyos datos estaban recogidos en la Encuesta Nacional de Salud Española (ENSE) 2011-12. La ENSE realiza conjuntamente entre el Ministerio de Sanidad y el Instituto Nacional de Estadística en virtud del convenio de colaboración suscrito entre ambos organismos. Su objetivo general es proporcionar la información necesaria sobre la salud de la población para poder planificar y evaluar las actuaciones en materia sanitaria.

Aunque se realiza de forma quinquenal, se eligió la ENSE del periodo 2011-12 frente a la realizada en el año 2017 por recoger aquella la percepción de la CVRS mediante el cuestionario EQ-5D-5L. Para asegurarse la representatividad de la ENSE, la selección de los sujetos se hace mediante un muestreo trietápico estratificado. Las unidades de primera etapa son las secciones censales, que se agrupan en estratos de acuerdo con el tamaño del municipio al que pertenecen las secciones. Las unidades de segunda etapa son las viviendas familiares principales. Dentro de cada hogar se selecciona a un adulto (15 o más años) para cumplimentar el cuestionario de adultos y en caso de que hubiera menores (de 0 a 14 años) se selecciona asimismo un menor para cumplimentar el cuestionario de menores. En la ENSE realizada en el año 2011-12 se seleccionó una muestra de aproximadamente 24.000 viviendas, 12 en cada una de las 2.000 secciones censales elegidas. (335). La encuesta para adultos fue respondida por 21007 personas y los microdatos están disponibles de manera pública. Se seleccionaron a los 6234 sujetos con edad igual o superior a 50 años que no referían tener diagnóstico de artrosis.

5.5. Variables

Se van a describir las variables utilizadas y su papel en el estudio.

5.5.1. *Variables de resultado*

Para dar respuesta a los tres primeros objetivos específicos la variable de resultado utilizada fue la transformación en utilidades del cuestionario genérico para la medida de la CVRS, EQ-5D (168). Como se ha señalado previamente, en su versión más actual, el cuestionario EQ-5D-5L (171) consta de dos partes: un termómetro para valorar la CVRS de 0 a 100, (escala visual analógica, EVA), y un cuestionario formado por 5 preguntas o dimensiones (movilidad, auto-cuidado,

realización de actividades habituales, dolor/malestar, y ansiedad/depresión) con 5 niveles de respuesta (1 sin problemas - 5 problemas extremos o imposibilidad). A partir de estas cinco preguntas se obtienen 3125 estados de salud, con una puntuación asociada a cada estado, denominada índice de utilidad. Inicialmente se consideraba que el índice de utilidad debía oscilar entre 0, estado que equivalía a la muerte y 1 que representaba la salud perfecta. Pero en los estudios de preferencias de los diferentes estados de salud se han encontrado estados menos preferibles que la muerte para la población general, lo que hace que este índice pueda tomar valores negativos, hasta -0,421 al utilizar los algoritmos publicados para España (190). La fiabilidad, validez y sensibilidad del EQ-5D-5L ha sido estudiada en pacientes con artrosis de miembro inferior en nuestro medio (336).

Ya se puso de manifiesto la relevancia de esta herramienta. Este cuestionario es el más ampliamente utilizado en el análisis de coste-efectividad de las tecnologías de atención de la salud (192). Las más prestigiosas organizaciones de evaluación de tecnologías sanitarias europeas, como el National Institute for Health and Care Excellence (NICE) en el Reino Unido, solicitan específicamente que el EQ-5D se utilice en todas las evaluaciones económicas que se les presenten para asegurar la comparabilidad entre los estudios (193,194). Las puntuaciones o “utilidades” derivadas de la última versión de esta herramienta, el EQ-5D-5L, han sido propuestas para informar los resultados de las evaluaciones económicas de tecnologías en España (190).

La variable de resultado para los objetivos específicos 4 y 5 fue la puntuación obtenida en la versión adaptada al español para España del cuestionario específico Oxford Hip Score (OHS). El OHS, es un cuestionario que puede ser cumplimentado en entrevistas “cara a cara”, de forma autoadministrada en el momento o ser remitido posteriormente a su cumplimentación. Consta de 12 preguntas con 5 respuestas cada una y evalúan la percepción de calidad de vida en las últimas 4 semanas del paciente (249).

La variable de resultado para los objetivos específicos 6 y 7 fue la puntuación obtenida en la versión adaptada al español para España del cuestionario específico Oxford Knee Score (OKS). El OKS es un cuestionario con similares características al OHS. Puede ser cumplimentado en entrevistas “cara a cara”, autoadministrado, o ser remitido posteriormente a su cumplimentación. Igualmente consta de 12 preguntas con 5 respuestas cada una y evalúan la percepción de calidad de vida en las últimas 4 semanas del paciente (296).

El OHS y el OKS, cuya descripción ya se recogía en el apartado 1.8 de la Introducción, fueron diseñados para estudiar los cambios en la CVRS tras procedimientos de remplazo total articular

(249,296), pero también se han utilizado tanto para valorar la situación basal de pacientes con artrosis de cadera o rodilla, o los cambios tras otros procedimientos quirúrgicos o no quirúrgicos (250,256,337).

Se ha comprobado que la administración telefónica tanto del OHS como del OKS produce resultados equivalentes a la utilización como cuestionario autoadministrado seguido de envío postal (338).

Para ambos cuestionarios, inicialmente, se utilizó una valoración distinta a la actual. Cada pregunta se puntuaba del 1 al 5, representando el 1 el mejor resultado/menos síntomas. Las puntuaciones de cada pregunta se sumaron para que la puntuación global fuera de 12 a 60 (siendo 12 el mejor resultado) (249,296). Algunos profesionales consideraron que esta puntuación no era intuitiva, por lo que adaptaron la puntuación, lo que provocó una gran confusión. De ahí la recomendación actual, de utilizar el 0 para el peor resultado y 4 el mejor, lo que permite un rango de respuestas entre 0 y 48, representando 48 el mejor resultado. (250). Cada una de las 12 preguntas de los dos índices están dispuestas de forma similar, con las categorías de respuesta que denotan menos (o ningún) síntoma a la izquierda de la escala (puntuación 4) y las que representan la mayor gravedad a la derecha (puntuación 0). En este estudio se ha utilizado este último formato de puntuación. La puntuación final permite clasificar la situación clínica en los sujetos con artrosis de cadera o rodilla en 4 grupos: excelente (>41puntos), buena (entre 34 y 41 puntos), moderada (entre 27 y 34 puntos) y mala (<27 puntos) (250).

La versión adaptada al castellano de ambos cuestionarios, OHS y OKS, fue obtenida bajo acuerdo con Oxford University Innovation TM, donde se produjo tras un proceso de traducción y retrotraducción inversa (ver Anexos 1 y 2).

5.5.2. Variables independientes.

Se incluyeron variables sociodemográficas, clínicas y sobre la percepción del estado de salud de los pacientes.

Como variables sociodemográficas se recogieron la edad y el sexo de los pacientes, si vivían solos o bajo cualquier forma de convivencia estable y el grupo social. El grupo social se categorizó según la clasificación de clase social ocupacional neoweberiana propuesta por Domingo-Salvany y cols en los siguientes grupos (339):

Caracterización de la Calidad de Vida Relacionada con la Salud en pacientes con artrosis de miembro inferior y validación de nuevas herramientas para su evaluación.

- Grupo I: directores/as y gerentes de establecimientos de 10 o más asalariados/as y profesionales tradicionalmente asociados a licenciaturas universitarias.
- Grupo II: directores/as y gerentes de establecimientos de menos de 10 asalariados/as, profesionales tradicionalmente asociados/as a diplomaturas universitarias y otros/as profesionales de apoyo técnico, deportistas y artistas.
- Grupo III: ocupaciones intermedias, asalariados/as de tipo administrativo y profesionales de apoyo a la gestión administrativa y de otros servicios, y trabajadores por cuenta propia.
- Grupo IV: supervisores/as y trabajadores/as en ocupaciones técnicas cualificadas.
- Grupo V: trabajadores/as cualificados/as del sector primario y otros/as trabajadores/as semicualificados/as.
- Grupo VI: trabajadores/as no cualificados/as

Entre las variables clínicas se recogieron:

- Fecha de diagnóstico de la artrosis, definida como se ha señalado previamente.
- Localización y lateralidad de la artrosis.
- Coexistencia de artrosis en otras articulaciones.
- Lugar de inclusión del paciente: centro hospitalario o centro de salud.
- Índice de masa corporal de los pacientes (IMC), recogido como variable continua y categorizado en: “bajo peso” ($IMC < 18,5$), “normopeso” ($18,5 \leq IMC < 25$), “sobrepeso” ($25 \leq IMC < 30$), “obesidad” ($IMC \geq 30$).
- Existencia de otras condiciones crónicas, definidas como cualquier situación de salud que requiere atención sanitaria y que se prolonga por un periodo superior a seis meses.
- Índice de Charlson. Este índice es un sistema de evaluación de la esperanza de vida a los diez años, en dependencia de la edad en que se evalúa, y de las comorbilidades del sujeto. Consta de 19 ítems, que influyen de una forma concreta en la esperanza de vida del sujeto. Inicialmente adaptado para evaluar la supervivencia al año, se adaptó finalmente en su forma definitiva para supervivencia a los 10 años (340). Este índice se ha utilizado

también para predecir el consumo de recursos de enfermos crónicos en el ámbito comunitario (341).

- Existencia de cirugías de remplazo articular previas de cadera o rodilla. Esta pregunta se repitió en el seguimiento a los seis meses.

Para caracterizar el estado de salud, además de los cuestionarios EQ-5D-5L, OHS y OKS descritos previamente se utilizaron las siguientes herramientas:

- Se recogió la percepción del estado de salud en una escala Likert de cinco categorías: muy bueno, bueno, regular, malo y muy malo.
- Para caracterizar el estadio clínico de la artrosis de miembro inferior se utilizó, además del OHS o el OKS, el Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) (198). El WOMAC es una escala multidimensional compuesta por 24 ítems que mide las dimensiones: dolor (5 ítems), rigidez (2 ítems) y función física (17 ítems) en pacientes con OAMI. Se utilizó la versión con cinco niveles de respuesta para cada ítem, que representan diferentes grados de intensidad (ninguno, leve, moderada, grave o extrema) que se califica de 0 a 4. Esa puntuación se suma y se estandariza de 0 a 100 (de mejor a peor capacidad). Cuanto mayor sea la puntuación, peor es el estado del paciente. Este cuestionario ha sido adaptado y validado en nuestro país (342).

5.6. Fuentes de información

La información relativa a los pacientes se obtuvo directamente de los mismos mediante los siguientes procedimientos. Una vez en la consulta el médico investigador le explicaba el objeto del estudio y pedía la firma del consentimiento informado. Seguidamente se recogía la información clínica pertinente en un cuaderno de recogida de datos y, o bien pasaba los cuestionarios sobre CVRS personalmente, o hacía pasar al paciente a otra sala donde personal entrenado previamente recogía dicha información cara a cara, o bien le decía al interesado que se llevase a su domicilios dichos cuestionarios en un sobre prefranqueado para ser devueltos. Se hacía seguimiento telefónico para intentar obtener la mayor tasa de respuesta posible.

Para evaluar la fiabilidad test-retest todos los pacientes incluidos en centros de la Comunidad de Madrid eran contactados telefónicamente entre 7 y 14 días después de la inclusión para volver a cumplimentar los cuestionarios OHS y OKS.

A los seis meses se programaba una nueva visita de seguimiento en la que el médico investigador reevaluaba la situación clínica del paciente y se le pedía que cumplimentase los cuestionarios sobre CVRS o bien presencialmente o en un momento posterior (se adjuntaba también sobre prefranqueado) y se hacía seguimiento telefónico de la respuesta.

Los datos de la ENSE fueron descargados de la página web del Ministerio de Sanidad donde están disponibles públicamente².

5.7. Análisis

En este apartado se describirán tanto las herramientas de análisis genéricas como las específicas para conseguir cada uno de los objetivos del estudio.

5.7.1. *Análisis descriptivo*

Se revisó la calidad de los datos haciendo un análisis exploratorio previo de frecuencias, y comprobando aquellos variables que ofrecían valores extremos o poco plausibles.

Se realizó el estudio descriptivo de cada variable cualitativa con frecuencia y porcentajes y con media y mediana, con sus correspondientes desviación estándar y rango intercuartílico, para las variables cuantitativas.

² Estos datos pueden descargarse libremente en el sitio web:
<https://www.mscbs.gob.es/estadisticas/microdatos.do>

5.7.2. Análisis inferencial

Modelos explicativos

Para dar respuesta a los objetivos específicos 1 a 3 se construyeron Modelos Lineales Generalizados (MLG, Generalized linear models, GLM), cuya variable dependiente fue el índice de utilidad atribuible al estado de salud descrito por el sujeto. Se eligió esta herramienta de análisis por su mayor tolerancia al incumplimiento de las asunciones necesarias para construir modelos clásicos, pues que permite obtener estimadores insesgados de las asociaciones en presencia de heterocedasticidad (343).

El MLG puede describirse de esta manera:

$$g\{E(y)\} = x_{ij}\beta_j \quad y \sim F$$

Donde y es la variable dependiente (la utilidad atribuida al estado de salud descrito por el paciente), $g()$ es la denominada función de enlace ("link function"), F es la familia de la distribución, x_{ij} representa la matriz de covariables y β_j el vector de los coeficientes.

La función de enlace relaciona la media condicional a las covariables. En este caso se seleccionó la función identidad como función de enlace $g()$, definida como

$$\eta = (\mu) = \mu$$

La familia de la distribución F , se utiliza para especificar la relación entre la varianza y la media. El paquete estadístico utiliza para la estimación de parámetros en el MLG los métodos de máxima verosimilitud, lo que nos permite obtener resultados sin tener que suavizar la variable dependiente y sin que la posible heterocedasticidad sea un problema (344,345). En cualquier caso y para evitar errores en la especificación del modelo se calcularon los errores de las estimaciones por métodos robustos (191,346).

Para seleccionar tanto el mejor modelo, como la mejor familia distribucional se estudiaron los criterios de información de Akaike y Bayes (Akaike Information Criteria, -AIC-, Bayes Information Criteria -BIC-).

La comparación entre modelos, si estos están anidados, mediante el AIC se hace de la siguiente manera (347). El mejor modelo es aquel que tiene el AIC con el valor mínimo, denotado por AIC*. El modelo candidato a sustituirle (por ejemplo porque sea más parsimonioso, esto es, que tenga menos variables explicativas) tendrá un AIC' y se denomina ΔAIC para el modelo candidato, denominado a la diferencia entre AIC' y AIC*.

La valoración de ΔAIC se puede hacer así:

- si es menor que 2, esto indica que hay evidencia sustancial para apoyar el modelo candidato (es decir, el modelo candidato es casi tan bueno como el mejor modelo).
- si está entre 4 y 7, indica que el modelo candidato tiene un apoyo considerablemente menor.
- si es superior a 10, indica que el modelo candidato no tiene prácticamente ningún apoyo (es decir, es poco probable que sea el mejor modelo).

Otra forma más compleja de interpretar la bondad de ajuste del modelo mediante los AIC es mediante el cálculo de los “pesos de Akaike” (“Akaike weights, A_w). Los A_w se obtienen normalizando los valores de probabilidad relativa. Es decir, son los ratios del AIC de un modelo candidato en relación con la suma de los AIC de todos los modelos candidatos, calculados de la siguiente manera.

$$A_w = \frac{e^{-0,5\Delta_m}}{\sum_{j=1}^M e^{-0,5\Delta_j}}$$

siendo Δ_m el ΔAIC del modelo que se está testando.

Los pesos de Akaike se interpretan como la probabilidad de que el modelo candidato sea el mejor entre el conjunto de modelos candidatos. Por ejemplo, si un modelo candidato tiene un peso de Akaike de 0,60, esto significa que, dados los datos, el modelo candidato tiene un 60% de probabilidad de ser el mejor (347).

También se estudió la mejora del BIC (348) de acuerdo a la siguiente interpretación. El mejor modelo es el que proporciona el mínimo BIC, denotado por BIC^* .

Podemos calcular el $\Delta BIC = BIC' - BIC^*$. Si tenemos m modelos, la magnitud del ΔBIC puede interpretarse como una prueba en contra de que un modelo candidato sea el mejor:

- menos de 2, el modelo candidato debe ser considerado.
- entre 2 y 6, la evidencia contra el modelo candidato es valorable.
- entre 6 y 10, la evidencia contra el modelo candidato es fuerte.
- Mayor de 10 el modelo candidato debe rechazarse.

MATERIAL Y MÉTODOS

Análisis de propiedades psicométricas.

Para dar respuesta a los objetivos específicos 4 a 7 se llevó a cabo un estudio de validación de un cuestionario evaluando los diferentes aspectos de: la factibilidad, la validez, la fiabilidad, la sensibilidad al cambio y la interpretabilidad.

La factibilidad se valoró a través de los sujetos que cumplimentan el cuestionario y de la preguntas que quedan sin contestar (314).

Dentro de los aspectos relativos a la validez se comenzaron evaluando los denominados “efecto suelo” y “efecto techo”. Se consideraba que estaban presentes cuando existía agrupamiento de la respuesta en la parte inferior (o superior) de la escala pues esto puede suponer un problema en el rango de respuesta ofrecido, limitando la capacidad de la herramienta para encontrar diferencias entre grupos (324). El límite de la agregación de respuestas para considerar la presencia de “efecto suelo” y/o “efecto techo” se estableció en el 15% de las respuestas, como recomienda la literatura (323).

Se analizó la validez convergente mediante la correlación de las escala del OHS y OKS con las escalas del WOMAC y con el EQ-5D-5L (índice de utilidad y EVA). Para estudiar la correlación se utilizaron la r de Pearson o la ρ de Spearman y se consideraron asociaciones fuertes todas aquellas por encima de 0,7, umbral utilizado para este fin (323). Para valorar la validez de grupos conocidos se compararon las puntuaciones de los instrumentos en evaluación (OHS y OKS) en los terciles de las distribuciones de los cuestionarios EQ-5D-5L y WOMAC respectivamente.

Respecto a la fiabilidad se valoraron la consistencia interna y la repetitividad o fiabilidad test-retest.

La consistencia interna se evaluó mediante análisis factorial y mediante el alfa de Cronbach. En el apartado introducción se señaló que el análisis de la consistencia interna correspondía a la valoración de la fiabilidad (322). Algunos autores consideran que algunas herramientas de valoración de la consistencia interna como el análisis factorial común, permiten explorar los constructos subyacentes, que no pueden medirse directamente, a través de ítems que se consideran medidas reflexivas del constructo, pues asumen que las puntuaciones individuales de los ítems son el resultado de un factor subyacente (349). En primer lugar se utilizó el análisis factorial exploratorio (AFE) para valorar la dimensionalidad de las escalas. Se comprobó la adecuación de la aplicación del análisis mediante el test de esfericidad de Bartlett y el test de

Kaiser-Meyer-Olkin (KMO). La hipótesis nula del test de Bartlett es que la matriz de correlaciones observadas es la matriz singular. El rechazo de esta hipótesis permite confirmar la existencia de relaciones lineales entre los factores y la variable explicada. Aunque es una prueba muy utilizada en la literatura, con muestras grandes tiene una utilidad limitada (350). El test de adecuación de la muestra de Kaiser-Meyer-Olkin es una medida de la varianza compartida por las variables y valores superiores a 0,90 se consideran óptimos (349). Se recogieron los pesos factoriales (considerando óptimos valores por encima de 0,40) y las comunalidades, que expresan el porcentaje de la varianza del ítem explicado por el factor.

Para complementar la valoración de la dimensionalidad de los cuestionarios, se realizó también un análisis factorial confirmatorio (AFC) para variables categóricas. Se utilizó el estimador robusto de mínimos cuadrados no ponderados y se calcularon varios índices de ajuste (351,352): el error cuadrático medio de aproximación (root mean square error of approximation, RMSEA), para el que se consideró aceptable un valor $<0,08$; y el índice de Tucker-Lewis (TLI) y el índice de ajuste comparativo (Comparative Fit Index, CFI), que debían ser ambos $>0,95$ (353). También se examinaron las cargas factoriales, y se consideraron aceptables las que eran $\geq 0,40$. Por lo tanto, si el modelo superaba estos criterios, se consideraba aceptable.

El alfa de Cronbach (354) se evaluó en las puntuaciones de la visita de inclusión. Este resume las correlaciones internas de todos los elementos de una escala. Cuanto mayor sea el coeficiente alfa (rango de 0,0 a 1,0), más consistente internamente es la escala y mayor será la probabilidad de que se está explotando una sola dimensión subyacente en el cuestionario. Para una herramienta unidimensional de 12 componentes se esperan valores del alfa de Cronbach superiores a 0,85 para que se considere óptima la consistencia interna (355).

La fiabilidad test-retest se estudió en la submuestra incluida en Madrid, comparando las puntuaciones del test con el retest (día 7 a 14) mediante el coeficiente de correlación intraclase (CCI, acuerdo absoluto). Siguiendo la clasificación propuesta para otras medidas de fiabilidad (356), valores del CCI superiores a 0,7 se consideran aceptables y por encima de 0,9 óptimos.

La sensibilidad al cambio se valoró a los seis meses de seguimiento. Para evaluar la capacidad de respuesta a los cambios que se producen en la evolución de la enfermedad, los instrumentos que se pretendían validar se volvieron a utilizar a los 6 meses de la inclusión de los pacientes. Se realizaron preguntas “transicionales” para valorar el cambio en el estado de salud percibido en relación con la artrosis respecto a los seis meses anteriores, que se contestaban en una escala de 5 respuestas (mucho peor, algo peor, igual, algo mejor o mucho mejor que antes). Estas preguntas

MATERIAL Y MÉTODOS

se utilizaron para valorar la sensibilidad al cambio tanto del OHS como del OKS. Las preguntas transicionales para el WOMAC se contestaban en la misma escala pero eran específicas para cada uno de sus dominios (dolor, rigidez y limitación). Los cambios en el OHS/OKS se estimaron restando las puntuaciones iniciales de las finales, con lo que los cambios positivos indican mejora del estado general. Para los cambios detectados con el EQ-5D se procedió de igual manera, pero los cambios en el WOMAC se estimaron restando las puntuaciones finales de las iniciales, de manera que en este caso también los resultados positivos indican mejoras. Se compararon las puntuaciones basales con las de seguimiento a los 6 meses, para ver si había habido cambios significativos, en cada grupo de pacientes según la pregunta transicional. Se calculó el tamaño del efecto como la relación entre la media y la desviación estándar del cambio para cada uno de los cambios, considerando que este era moderado para valores superiores a 0,5 y grande para valores superiores a 0,8 (357), y se comparó con el encontrado para las escalas del WOMAC y del EQ-5D-5L.

Además, estimamos la mínima diferencia clínicamente importante (MDCI) y el mínimo cambio detectable (MCD). Estas dos medidas están relacionadas con la capacidad de respuesta, pero están más orientadas a la clínica y centradas en el nivel individual. La MDCI se estimó con el cambio medio en los pacientes que habían notado mejoría moderada (se encuentran “algo mejor”) a los 6 meses (358).

El MCD expresa la magnitud mínima de cambio por encima de la cual es probable que el cambio observado sea real y no refleje solo la variabilidad de la medición. Para la estimación del MCD, es necesario calcular previamente el error estándar de medición (EEM o Standard Error of Measure, SEM), que representa la cantidad de error asociada a la evaluación de un sujeto individual. Este se estimó como la raíz cuadrada del término de error cuadrático medio del ANOVA (359,360). A partir del EEM, el MCD se derivó de la siguiente manera (323,359):

$$\text{MCD} = \text{EEM} \times Z_{\text{score}} \times \sqrt{2}$$

Se estableció un nivel de confianza del 95% (MCD 95%), correspondiente a un valor Z score de 1,96. La interpretación del MCD 95% es el umbral por encima del cuál un cambio detectado en un paciente puede ser valorado como un cambio real y no solo resultado de un error de medición con una confianza del 95%.

Para valorar la interpretabilidad debe relacionarse el MDCI con MCD95%. Si la relación entre ambos es superior a 1 se puede afirmar que la MDCI puede discriminarse del error de medición (325). El MCD 95% puede calcularse a nivel individual y para el grupo de sujetos.

El nivel de significación estadística se estableció en 0,05 y el programa estadístico utilizado fue Stata 14[®].

5.8. Aspectos éticos y legales

El estudio respetó los principios éticos recogidos en la Declaración de Helsinki y sus revisiones.

Todos los pacientes incluidos dieron previamente su consentimiento por escrito a participar en el estudio.

Se preservó la confidencialidad de la información en cumplimiento de la Ley Orgánica 15/1999 de 13 de diciembre de protección de datos de carácter personal. Tras extraer la información y codificarla en variables se elaboró una base de datos disociada, de manera que solo tenían acceso a los datos identificativos los investigadores principales del estudio.

Todos los datos del trabajo de campo fueron recogidos por los investigadores y colaboradores de los proyectos de investigación PI PI1300560, PI1300518 y PI1300648 (financiados por el Instituto de Salud Carlos III y el Fondo Europeo de Desarrollo Regional (FEDER), “Una manera de hacer Europa”).

Este estudio recibió la aprobación de los siguientes comités de ética para la investigación clínica: CEIC de Euskadi (PI2014050), Hospital Fundación Jiménez Díaz (PIC 80/2013_HRJC), Hospital Universitario de Fuenlabrada (APR 14-27), Hospital Universitario Fundación Alcorcón (14/19), Hospital Universitario de Canarias (2014-109), Hospital Universitario Nuestra Señora de Candelaria (PI-09/15).

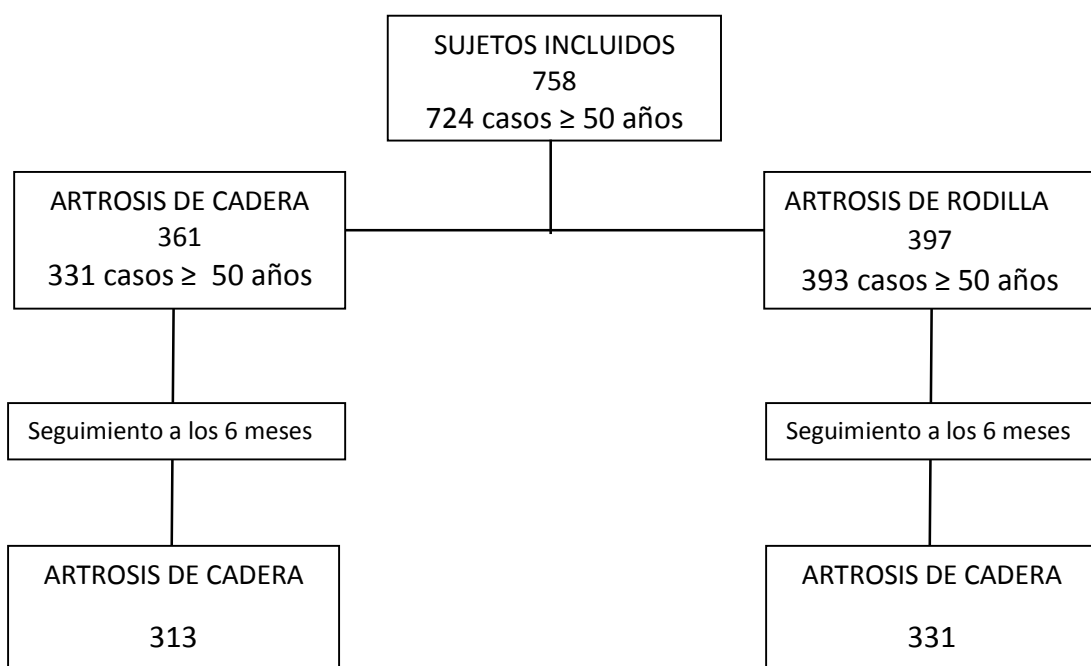
6. Resultados

RESULTADOS

Fueron incluidos en el estudio 758 sujetos con artrosis de cadera o rodilla (ver figura 1). De estos 361 fueron seleccionados por tener diagnosticada artrosis de cadera y 397 por artrosis de rodilla. En ese momento, 724 sujetos tenían 50 o más años y sus datos fueron utilizados en el estudio para medir el impacto de la artrosis de miembro inferior en la CVRS, 331 con artrosis de cadera y 393 con artrosis de rodilla.

De los 361 sujetos incluidos inicialmente con artrosis de cadera, a los 6 meses completaron el seguimiento 313 sujetos. De los 397 sujetos incluidos con artrosis de rodilla, 331 completaron el seguimiento a los 6 meses.

Figura 1. Diagrama de flujo de la población incluida



De los 758 sujetos incluidos, 315 (41,6%) lo fueron en Vizcaya, 282 (37,2%) en Madrid y 161 (21,2%) en Canarias. Del total de pacientes, 404 (53,3%) fueron reclutados en consultas de Traumatología, 295 (38,9%) en consultas de Atención Primaria, y 59 (7,8%) en consultas de Reumatología.

RESULTADOS

En la tabla 1 se muestran las características de la población incluida

Tabla 1. Características de la población incluida

Características		Artrosis de cadera				Artrosis de rodilla			
		N	%	Media	DS	N	%	Media	DS
Edad		361		67.88	11.67	397		71.42	9.06
Sexo	Mujer	192	53,2			277	69,8		
	Hombre	169	46,8			120	30,2		
Tipo de centro de salud	Hospital	212	58,7			251	63,2		
	Atención Primaria	149	41,3			146	36,8		
Región	País Vasco	157	43,5			158	39,8		
	Canarias	80	22,2			81	20,4		
	Madrid	124	34,3			158	39,8		
Clase social	Grupo I	26	7,2			16	4,0		
	Grupo II	14	3,9			14	3,5		
	Grupo III	47	13,0			47	11,8		
	Grupo IV	59	16,3			65	16,4		
	Grupo V	81	22,4			90	22,7		
	Grupo VI	53	14,7			67	16,9		
	No consta	81	22,4			98	24,7		
Articulación con artrosis	Derecha	138	38,2			110	27,7		
	Izquierda	121	33,5			119	30		
	Ambas arts.	102	28,2			168	42,3		
Artrosis en otras articulaciones	Cadera					53	13,3		
	Rodilla	115	31,8						
	Otras					228	57,4		
Artroplastia contralateral	Cadera o Rodilla	63	17,4			78	19,6		
Tiempo desde el diagnóstico (años)	<1	118	32,7			62	15,6		
	1 a <5	152	42,1			181	45,6		
	5 a <10	63	17,4			93	23,4		
	10 a <15	14	3,9			35	8,8		
	≥15	8	2,2			21	5,3		
	No consta	6	1,7			5	1,3		
Índice de masa corporal		355		28.16	4.57	389		29.67	4.87
Índice de Charlson (comorbilidades)		361		0.84	1.31	395		0.78	1.14
DS: Desviación estandar									

Dentro de los 361 sujetos incluidos con artrosis de cadera, predominan las mujeres (53,2%) frente a los hombres (46,8%), pertenecen, en general, a grupos sociales poco desfavorecidos. El 22,4% de los pacientes desarrollan actividades cualificadas en el sector primario o son trabajadores semicualificados, y hasta el 14,7 % desarrollan actividades no cualificadas. Destaca el elevado índice de masa corporal de la muestra, con una media de IMC superior a 28, lo que supone una situación de sobrepeso.

Encontramos que hasta el 28,2% de los casos presentan afectación de ambas caderas, y el 31,8% presenta además, afectación de la rodilla. El tiempo transcurrido desde el diagnóstico de coxartrosis hasta la inclusión en el estudio, es, en un 75% de los casos, inferior a 5 años y el 17,4% de los pacientes presentaban una intervención de reemplazo articular en la cadera contralateral.

En cuanto a los pacientes con artrosis de rodilla, se incluyeron un total de 397 sujetos, con predominancia de las mujeres de forma más marcada que el caso de la artrosis de cadera; representan casi el 70% de la muestra frente al 30% de hombres. Al igual que ocurre con la cadera, los pacientes con gonartrosis, pertenecían a los grupos sociales que desarrollan actividades no cualificadas o son trabajadores del sector primario, representando casi el 40% entre ambos grupos pertenecientes a los grupos V y VI de la clasificación de clase social ocupacional neoweberiana. También destaca el elevado índice de masa corporal que presenta el grupo, con un IMC de media superior a 29, lo que supone una situación de sobrepeso, muy cercana a la obesidad.

Al igual que en la cadera, encontramos una alta incidencia de afectación de artrosis en otras articulaciones; estando presente hasta en el 70% de los casos, de los cuales el 13% corresponden a afectación en forma de coxartrosis. Además, cerca del 20% de pacientes habían sido sometidos a cirugía de reemplazo articular en la rodilla contralateral.

Respecto al tiempo transcurrido desde el diagnóstico de gonartrosis hasta la inclusión en el estudio, este es mayor que el descrito para el caso de la artrosis de cadera, el 61% tenían un diagnóstico de menos de 5 años de antigüedad, de los cuales solo 15% llevaba menos de 1 año desde diagnóstico, frente al 32% que encontramos con la artrosis de cadera.

El estado de salud percibido y la utilidad expresada por los pacientes con artrosis de cadera y rodilla se muestran en la tabla 2. Los pacientes con artrosis de cadera presentan un deterioro en la CVRS que se muestra representado por una puntuación media en el cuestionario EQ-5D-5L en término de transformación en utilidades de 0,520 (donde 0 es el estado de salud correspondiente a la muerte y 1 representa la salud perfecta). El valor medio del EQ-5D en la escala analógica visual es de 54,47. En cuanto a los instrumentos específicos para medir la CVRS en los pacientes con

RESULTADOS

artrosis de cadera, el cuestionario WOMAC muestra una merma de la CVRS en las tres dimensiones (dolor, rigidez y función) con una puntuación media global de 50,69. El cuestionario de Oxford para la cadera muestra una puntuación media de 22,84 (siendo el 0 el peor estado de salud y 48 el mejor estado de salud posible), con una distribución de la muestra que clasifica al 65% de los pacientes como de una situación clínica mala, es decir, que registran puntuaciones en la escala de Oxford inferiores a 27.

Los resultados en los pacientes con artrosis de rodilla, también presentan una merma importante en la CVRS con una puntuación media en el cuestionario EQ-5D-5L expresados en utilidades de 0,544. El valor medio del EQ-5D en la escala analógica visual es de 57,36. Respecto a los cuestionarios específicos, el WOMAC muestra una puntuación media global de 49,63, con afectación de las 3 dimensiones del cuestionario. El cuestionario Oxford de rodilla muestra una puntuación media de 21,96, con una distribución que clasifica al 69% de los pacientes en una situación clínica mala, resultados similares a los obtenidos para la cadera.

Tabla 2. Percepción del estado de salud en los pacientes con artrosis de cadera y rodilla.

Características		Artrosis de cadera				Artrosis de rodilla			
		N	%	Media	DS	N	%	Media	DS
WOMAC puntuación de dolor (0-100)		360		45,81	22,77	396		47,00	20,49
WOMAC – Rigidez (0-100)		360		48,26	26,05	396		46,81	25,42
WOMAC – Funcionalidad física (0-100)		360		52,43	23,05	397		51,08	20,95
Puntuación WOMAC (0-100)		359		50,69	22,25	397		49,63	20,32
Cuestionario Oxford (0-48)		359		22,84	10,62	395		21,96	9,96
EQ-5D-5L Utilidades		357		0,520	0,304	393		0,544	0,271
EQ-5D EVA (0-100)		359		54,47	22,27	390		57,36	21,55
Oxford Score	Excelente	17	4,7			13	3,3		
	Bueno	49	13,6			37	9,4		
	Moderado	61	17,0			73	18,5		
	Malo	232	64,6			272	68,8		
DS: Desviación estandar; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index. WOMAC: 0, mejor ; 100 peor EQ-5D-5L Utilidades: 0, muerte ; 1, perfecto estado de salud Puntuación Cuestionario de Oxford : 0, peor; 48 mejor estado posible									

6.1. Estudio del impacto de la artrosis de miembro inferior en la CVRS

Para estudiar el impacto de la artrosis de miembro inferior, se seleccionaron de entre los sujetos incluidos, aquellos mayores de 50 años, y de los sujetos incluidos en la ENSE 2011-12 aquellos mayores de 50 años que no reportaban diagnóstico de artrosis.

Se presentan los resultados para 6958 sujetos mayores de 50 años, 6234 provenientes de la población general sin artrosis de miembros inferiores y 724 (10,4% del total) con diagnóstico de OAMI. De los pacientes diagnosticados de OAMI, 393 (54,3%) presentaban diagnóstico de artrosis de rodilla y 331 (45,7%) de cadera. El reclutamiento de los pacientes con OAMI se produjo para 286 en consultas de Atención Primaria (39,5%), y para 438 (60,5%) en Atención Especializada, 381 (52,6%) en consultas de Traumatología y 57 (7,9%) en consultas de Reumatología.

En la tabla 3 se presentan las características del total de la muestra estudiada. Entre los pacientes con OAMI predominan las mujeres (62%), con un aumento de la prevalencia que se incrementa con la edad; desde los 50 hasta los 80 años, que se ve limitado en las edades extremas, y con un pico de prevalencia en la séptima década de la vida, que representa el 38% del total. Predominan los sujetos pertenecen a los grupos sociales menos favorecidos, más de la mitad de la muestra estaba clasificada en los grupos V y VI de la clasificación de clase social ocupacional neoweberiana. También hubo mayor prevalencia en las personas que conviven con pareja, y destaca el elevado índice de masa corporal de los pacientes de la muestra, que sitúa a más del 80% en una situación de sobrepeso u obesidad.

Tabla 3. Características de la muestra estudiada

	Población general N=6234	Población con OAMI N=724	p
Edad (%)			
50- 54 años	21,82	4,70	<0.001
55-59 años	18,53	8,15	
60-64 años	15,77	13,12	
65-69 años	13,97	16,57	
70-74 años	9,62	17,54	
75-79 años	8,87	20,99	
80-84 años	6,26	13,40	
≥ 85 años	5,17	5,52	
Sexo (%)			
Mujer	47,06	62,71	<0.001
Clase social (%)			
Grupo I	11,64	7,12	<0.001
Grupo II	7,93	4,56	
Grupo III	19,00	16,79	
Grupo IV	15,22	20,62	
Grupo V	32,02	29,38	
Grupo VI	14,20	21,53	
Convivencia (%)			
Pareja	61,93	66,99	<0.001
Enfermedades crónicas (%)			
Ninguna	17,85	0,00	<0.001
Una	22,44	55,11	
Dos	20,82	30,39	
Tres o más	38,88	14,50	
IMC (%)			
Bajo peso	1,03	0,14	<0.001
Normopeso	31,18	19,06	
Sobrepeso	40,02	43,23	
Obesidad	27,77	37,57	
Estado de salud percibido (%)			
Muy bueno	13,60	2,23	<0.001
Bueno	55,25	22,67	
Regular	23,24	48,26	
Malo	6,42	21,56	
Muy malo	1,49	5,29	
EQ-5D-5L, utilities, media (DS)	0,924 (0,160)	0,532 (0,287)	<0,001
EQ-5D-5L, EVA, media (DS)	74,71 (17,71)	56,31 (21,70)	<0,001

La percepción del estado de salud y la CVRS expresada por los pacientes con OAMI fue significativamente peor que para la población general, tanto para las transformaciones del EQ-5D-5L en utilidades, como para su EVA (tabla 3), como para cada una de las dimensiones del cuestionario (tabla 4). Los pacientes con OAMI muestran una puntuación en EQ-5D-5L en términos de utilidad de 0,532, frente al 0,924 de la población general, lo que supone una pérdida de utilidad de 0,392 puntos. Cerca del 30% de los pacientes con OAMI describen un estado de salud malo o muy malo, frente al 8% de la población general.

Encontramos diferencias significativas en la CVRS en las 5 dimensiones del cuestionario EQ-5D-5L, siendo más acusadas en los dominios de movilidad, realización de actividades habituales y dolor/malestar. En la dimensión de movilidad el 43% de los pacientes responden tener problemas moderados, frente al 4% de la población general, y solo el 10% refiere no tener problemas de movilidad frente al 84% de la población general. Para la realización de actividades de la vida diaria más del 50% de los pacientes refieren tener problemas moderados o graves, frente al 4% de la población general. Y en la dimensión del dolor o malestar, cerca del 70% de los pacientes responden presentar dolor moderado o grave, frente a menos del 10% de la población general.

El rango de “utilidades” para los estados de salud expresados osciló entre -0,416 (estado 55555) y 1 (estado 11111), tanto para la población general como para los pacientes con OAMI. En el caso de los pacientes con artrosis de rodilla el peor estado de salud reportado fue 44555 al que se le atribuye una utilidad de -0,297.

RESULTADOS

Tabla 4. Distribución de las respuestas a los diferentes dominios del EQ-5D-5L en población general y con artrosis de miembro inferior (OAMI).

Dimensión EQ-5D-5L	Porcentaje de respuesta por cada nivel					p
	1	2	3	4	5	
Movilidad						
Población general n=6234	84,62	7,92	4,23	2,29	0,93	<0,001
OAMI n=723	10,93	20,06	43,02	23,93	2,07	
Autocuidado	1	2	3	4	5	
Población general n=6234	93,55	2,98	1,44	0,90	1,12	<0,001
OAMI n=722	28,39	27,70	31,58	11,36	0,97	
Actividades diarias						
Población general n=6234	89,09	5,21	2,68	1,57	1,44	<0,001
OAMI n=722	16,48	25,21	36,70	16,62	4,99	
Dolor/ disconfort						
Población general n=6231	76,46	14,35	6,55	2,33	0,32	<0,001
OAMI n=723	6,22	21,30	37,07	30,57	4,84	
Ansiedad/ Depresión						
Población general n=6226	85,21	9,14	3,82	1,48	0,35	<0,001
OAMI n=716	44,13	23,18	18,30	11,31	3,07	

La población con artrosis de miembro inferior incluida en este subestudio (mayores de 50 años) tiene una situación clínica casi superponible al total de la población estudiada, descrita anteriormente. En la tabla 5 se presenta el nivel de gravedad para los pacientes con artrosis de rodilla o cadera, medida con el WOMAC y con las escalas de Oxford, OHS/ OKS.

Los pacientes con artrosis de rodilla tenían afectación bilateral con más frecuencia (42%), que los pacientes con coxartrosis (28%). Sin embargo los pacientes con artrosis de cadera tenían afectación de rodilla concomitante en más casos, 35% frente al 13% de los pacientes con artrosis de rodilla y afectación de la cadera.

No había diferencias significativas entre la CVRS o el nivel de gravedad medido con los instrumentos específicos entre los pacientes con artrosis de rodilla o cadera. Solo en la EVA del EQ-5D-5L se apreciaba que los pacientes con artrosis de cadera podían tener una peor percepción que los diagnosticados de artrosis de rodilla aunque la diferencia quedaba al borde de la significación estadística.

RESULTADOS

Tabla 5. Características de los pacientes con OAMI incluidos en el estudio.

	Pacientes con OA de cadera n=331	Pacientes con OA de rodilla n=393	p
Afectación bilateral (%)	28,40	42,24	<0,001
Afectación de otra gran articulación de miembro inferior (cadera o rodilla) (%)	34,74	13,49	<0,001
EQ-5D-5L Utilidades. Media (DS)	0,517 (0,303)	0,543 (0,272)	0,110
EQ-5D-5L EVA. Media (DS)	54,92 (21,78)	57,47 (21,59)	0,059
Oxford score. Media (DS)	22,72 (10,58)	21,97 (9,94)	0,163
Puntuación WOMAC. Media (DS)			
Puntuación WOMAC dolor	45,82 (22,63)	47,09 (20,55)	0,216
Puntuación WOMAC limitación	52,65 (22,81)	51,14 (20,99)	0,178
Puntuación WOMAC rigidez	48,33 (25,97)	46,81 (25,29)	0,213
Puntuación WOMAC total	50,85 (22,02)	49,95 (20,10)	0,282
Estado según Oxford score (%)			
Excelente	4,56	3,32	0,348
Bueno	11,85	8,42%	
Moderado	18,24	19,64	
Malo	65,35	68,32	

La tabla 6 recoge los resultados de los modelos explicativos para las utilidades expresadas. Se construyeron 3 modelos. El modelo 1 incluye variables que se han relacionado con la CVRS expresada, el modelo 2 añade la presencia de OAMI y la articulación afectada y el modelo 3 incluye, además, si el paciente ha sido valorado en el entorno hospitalario o no. No se incluyó la variable “percepción del estado de salud” como variable explicativa por el riesgo de sobreajuste al medir el mismo constructo que la variable dependiente. Ni la bilateralidad, ni el tener afectado otro grupo articular por artrosis en miembro inferior (cadera o rodilla) mejoraban la capacidad explicativa del modelo final (modelo 3). El modelo 3 se presenta como el mejor modelo pues sus “Akaike weights” respecto al conjunto de modelos y respecto al modelo 2, tienen valor 1, esto es hay “certeza estadística” de que es el mejor modelo. Bajo el criterio de la mejora del BIC la evidencia contra la bondad de los modelos 1 y 2 es muy fuerte (diferencias de BIC muy por encima de 10 en ambos casos).

En el modelo 3 se observa como la presencia de artrosis de cadera disminuye las utilidades en personas de similar edad, mismo sexo, grupo social, estado de convivencia y similar carga de enfermedad crónica en una media de 0,347 puntos. Para los pacientes que presentan artrosis de rodilla la disminución de este valor es de 0,295 puntos de media. Para cualquier tipo de artrosis el hecho de ser atendido en el medio hospitalario se asocia con una disminución media de las utilidades de 0,112 puntos, ajustado por las características referidas.

Se asociaron a un peor nivel de “utilidad”, además de la OAMI y ser atendido en un entorno hospitalario, una mayor edad, pertenecer a un grupo social más desfavorecido, ser mujer, vivir solo, y presentar más condiciones crónicas.

Tabla 6. Modelos explicativos de las diferencias en las utilidades derivadas del EQ-5D-5L

	Modelo 1	Modelo 2	Modelo 3
Variable	Coef (IC 95%)	Coef (IC 95%)	Coef (IC 95%)
Nivel Atención			
Atención hospital vs resto			-0,112 (-0,158- -0,065)
Artrosis MI			
Cadera Vs No artrosis		-0,406 (-0,444- -0,369)	-0,347 (-0,390- -0,303)
Rodilla Vs No artrosis		-0,365 (-0,396- -0,334)	-0,295 (-0,336- -0,255)
Edad			
55-59 vs 50-54 años	-0,011 (-0,023- 0,001)	-0,004 (-0,014- 0,006)	-0,004 (-0,014- 0,006)
60-64 vs 50-54 años	-0,013 (-0,025- 0,000)	0,007 (-0,003- 0,017)	0,007 (-0,003- 0,018)
65-69 vs 50-54 años	-0,015 (-0,029- -0,001)	0,013 (0,002- 0,024)	0,013 (0,002- 0,024)
70-74 vs 50-54 años	-0,037 (-0,054- -0,021)	0,005 (-0,009- 0,019)	0,005 (-0,009- 0,019)
75-89 vs 50-54 años	-0,054 (-0,073- -0,036)	0,003 (-0,012- 0,018)	0,003 (-0,012- 0,018)
80-84 vs 50-54 años	-0,121 (-0,147- -0,096)	-0,062 (-0,084- -0,040)	-0,064 (-0,086- -0,042)
≥ 85 vs 50-54 años	-0,175 (-0,209- -0,141)	-0,148 (-0,182- -0,114)	-0,150 (-0,184- -0,117)
Sexo			
Varón vs mujer	0,042 (0,032- 0,051)	0,022 (0,014- 0,030)	0,022 (0,014- 0,030)
Clase social			
Grupo II vs grupo I	-0,005 (-0,021- 0,011)	-0,009 (-0,022- 0,005)	-0,010 (-0,024- 0,004)
Grupo III vs grupo I	-0,024 (-0,039- -0,009)	-0,021 (-0,034- -0,008)	-0,021 (-0,034- -0,008)
Grupo IV vs grupo I	-0,030 (-0,046- -0,014)	-0,017 (-0,030- -0,004)	-0,017 (-0,030- -0,004)
Grupo V vs grupo I	-0,023 (-0,037- -0,010)	-0,023 (-0,035- -0,012)	-0,023 (-0,035- -0,012)
Grupo VI vs grupo I	-0,042 (-0,059- -0,026)	-0,024 (-0,038- -0,010)	-0,024 (-0,038- -0,010)
Convivencia			
Pareja vs solos	-0,006 (-0,016- 0,004)	0,010 (0,002- 0,019)	0,010 (0,002- 0,019)
Enfermedades crónicas			
Una vs ninguna	-0,075 (-0,086- -0,064)	-0,012 (-0,020- -0,004)	-0,010 (-0,018- -0,003)
Dos vs ninguna	-0,058 (-0,069- -0,047)	-0,028 (-0,037- -0,019)	-0,030 (-0,039- -0,021)
Tres o más vs ninguna	-0,090 (-0,100- -0,081)	-0,094 (-0,102- -0,085)	-0,093 (-0,102- -0,084)
IMC			
Bajo peso vs normopeso	-0,037 (-0,098- 0,024)	-0,047 (-0,101- 0,007)	-0,047 (-0,100- 0,006)*
Sobrepeso vs normopeso	-0,008 (-0,018- 0,002)	0,004 (-0,004- 0,013)	0,004 (-0,005- 0,012)
Obesidad vs normopeso	-0,051 (-0,063- -0,038)	-0,028 (-0,039- -0,018)	-0,029 (-0,040- -0,019)
*0,05<p<0,1			
Características del modelo	N= 6542 AIC= -3360,13 BIC= -3217,62	N= 6542 AIC= -5529,77 BIC= -5373,69	N= 6542 AIC= -5593,02 BIC= -5430,15

Para dar respuesta a los objetivos 2 y 3 se volvieron a construir modelos explicativos pero sólo para los sujetos mayores de 50 años que tenían diagnóstico de artrosis de miembro inferior, incluidos en el estudio. Se tomó como base el modelo 3, el mejor modelo explicativo para la comparación descrita previamente incluyendo la bilateralidad y la presencia de artrosis en otra articulación de miembro inferior como variables explicativas. Este será el modelo 1'. En el modelo 2' se eliminaron variables sociales que no resultaban explicativas, en concreto la convivencia y el grupo social. En el modelo 3' se eliminó la variable presencia de artrosis en otra articulación de miembro inferior al no aportar capacidad explicativa al modelo (tabla 7).

La selección del mejor modelo mediante los "Akaike weights" no fue sencilla en este caso. Utilizando los AIC mostrados en la tabla 7, el modelo que presenta más probabilidad de ser el mejor es el 1'. Si se utilizasen los valores de AIC ofrecidos por Stata 14[®], que difieren en su método de cálculo de los clásicos (datos no presentados), el mejor modelo sería el 3' aunque con poca confianza, apenas un 2% de probabilidad de ser mejor que el 1'. Bajo los criterios de BIC, interpretando las recomendaciones de Kass y Raftery, el modelo 3' es netamente superior al 1' y razonablemente mejor que el 2'. Se eligió el modelo 3' como el mejor por las razones expuestas y por ser el más parsimonioso.

En el modelo 3' se observa como la presencia de artrosis de rodilla se asocia con puntuaciones de utilidad de 0,055 superiores a los pacientes con artrosis de cadera ajustado por edad, IMC, lugar de atención y existencia de artrosis bilateral. El lugar de atención es importante a la hora de valorar a los pacientes. Aquellos atendidos en hospitales tiene de media 0,139 puntos menos de utilidad que los atendidos en Atención Primaria. Ser varón se asocia a una mejor percepción de la CVRS medida a través de las utilidades (asociación al borde de la significación estadística) y tanto el bajo peso, como la obesidad se asocian a peor percepción de CVRS en pacientes con artrosis de miembro inferior.

El hecho de que la artrosis sea bilateral también puede asociarse con una peor percepción de CVRS en pacientes con artrosis de miembro inferior, pero dicha asociación no alcanzó el nivel de significación establecido como umbral en este estudio.

RESULTADOS

Tabla 7. Modelos explicativos de las diferencias en las utilidades derivadas del EQ-5D-5L solo para pacientes con artrosis de miembro inferior

	Modelo 1	Modelo 2	Modelo 3
Variable	Coef (IC 95%)	Coef (IC 95%)	Coef (IC 95%)
Nivel Atención			
Atención hospital vs A.P.	-0,109 (-0,158- -0,061)	0,133 (-0,174- -0,091)	-0,133 (-0,174- -0,091)
Artrosis MI			
Rodilla vs cadera	0,051 (0,000- 0,101)	0,056 (0,013- 0,100)	0,055 (0,013- 0,097)
Edad			
55-59 vs 50-54 años	0,000 (-0,134- 0,135)		
60-64 vs 50-54 años	0,047 (-0,0765- 0,170)		
65-69 vs 50-54 años	0,034 (-0,096- 0,163)		
70-74 vs 50-54 años	0,110 (-0,012- 0,232)		
75-89 vs 50-54 años	0,063 (-0,060- 0,186)		
80-84 vs 50-54 años	-0,003 (-0,140- 0,132)		
≥ 85 vs 50-54 años	0,075 (-0,098- 0,248)		
Sexo			
Varón vs mujer	0,054 (0,005- 0,103)	0,042 (-0,001- 0,085)	0,042 (-0,001- 0,084)*
Clase social			
Grupo II vs grupo I	0,035 (-0,090- 0,159)		
Grupo III vs grupo I	-0,0270 (-0,134- 0,080)		
Grupo IV vs grupo I	-0,050 (-0,156- 0,056)		
Grupo V vs grupo I	-0,029 (-0,130- 0,071)		
Grupo VI vs grupo I	-0,048 (-0,153- 0,057)		
Convivencia			
Pareja vs solos	-0,014 (-0,066- 0,038)		
Enfermedades crónicas			
Dos vs solo una*	-0,027 (-0,080- 0,025)	-0,066 (-0,112- -0,021)	-0,066 (-0,112- -0,021)
Tres o más vs una*	-0,108 (-0,183- -0,032)	-0,118 (-0,183- -0,054)	-0,118 (-0,183- -0,053)
IMC			
Bajo peso vs normopeso	0,587 (-0,680- -0,494)	-0,558 (0,623- -0,493)	-0,560 (-0,621- -0,499)
Sobrepeso vs normopeso	-0,044 (-0,103- 0,015)	-0,030 (-0,082- 0,022)	-0,030 (-0,083- 0,022)
Obesidad vs normopeso	-0,103 (-0,166- -0,040)	-0,089 (-0,145- -0,034)	-0,089 (-0,144- -0,034)
Bilateralidad			
	-0,028 (-0,077- 0,203)	-0,036 (-0,079- 0,006)	-0,036 (-0,078- 0,006)*
Otra artrosis en MMII	0,005(-0,057- 0,067)	0,006 (-0,048- 0,060)	
*0,05<p<0,1			
Características del modelo	N= 543 AIC= 155,04 BIC= 253,88	N= 716 AIC= 192,88 BIC= 238,62	N= 716 AIC=190,94 BIC= 232,10

6.2. Validación del OHS

Para validar el cuestionario OHS, fueron incluidos un total de 361 pacientes, 157 en Vizcaya, 124 en Madrid y 80 en Canarias (ver tabla 1). De estos un 41,3% fueron incluidos en consultas de Atención Primaria, un 47,9% en consultas de Traumatología y un 10,8% en consultas de Reumatología. El 53,2% de los pacientes eran mujeres (IC 95%: 48,0-58,4%) con una edad media para la muestra de 67,8 años (IC95%: 66,7-69,1 años).

Respecto a la localización de la artrosis, el 38,2% (IC95%: 33,2- 43,3%) tenían afectada la cadera derecha, el 33,5% (IC95%: 28,6- 38,4%) la izquierda y el 28,2% (IC95%: 23,6-32,92%) de ambas.

El 17,4% (IC 95%: 13,6- 21,4%) habían sido sometidos a cirugía de prótesis de cadera en la articulación contralateral previamente. El Índice de comorbilidad de Charlson tenía una media de 0,8 puntos (IC95%: 0,7-1,0). El IMC medio era de 28,2 (IC 95%: 27,7- 28,6).

En la tabla 8 se presentan los resultados expresados por los pacientes para los cuestionarios OHS, WOMAC y EQ-5D-5L.

Tabla 8. Puntuaciones expresadas en los cuestionarios OHS, WOMAC y EQ-5D-5L

	N	Media puntuaciones (IC 95%)	Mediana (Recorrido Intercuartil)
Cuestionario OHS	359	22,8 (21,7- 23,9)	22,0 (15,0-30,0)
WOMAC dolor	360	45,8 (43,4- 48,2)	45,0 (30,0- 60,0)
WOMAC rigidez	360	48,3 (45,6- 51,0)	50,0 (25,0- 62,5)
WOMAC función física	360	52,4 (50,0- 54,8)	51,5 (36,8- 70,6)
EQ-5D-5L utilidad	357	0,52 (0,49-0,55)	0,61 (0,32- 0,73)
EQ-5D-5L escala EVA	357	54,5 (52,2- 56,8)	50,0 (40,0-70,0)
OHS: Oxford Hip Score: 0 peor resultado; 48 mejor estado posible WOMAC: 0, mejor; 100 peor EQ-5D-5L Utilidades: 0, muerte ; 1, perfecto estado de salud EQ-5D EVA (0-100): 0, peor estado de salud y 100 mejor estado de salud imaginable			

6.2.1. *Aceptabilidad*

Se obtuvo información que permitía calcular el valor resumen del cuestionario OHS en 359 casos (99,4%; IC95%: 98,7-100%). Las preguntas 3 y 4 fueron contestadas en todos los casos, y las preguntas 2, 5, 7, 8, y 10 en todos los casos menos uno. Las preguntas 6 y 11 fueron dejadas sin contestar en dos ocasiones y la 9 y la 12 en tres cuestionarios.

6.2.2. *Validez de contenido*

“Efecto suelo” y “efecto techo”

En todas las preguntas se utilizaron todos los rangos de respuesta posible (de 0 a 4). En ninguna pregunta se acumularon más de un 35% de las respuestas en ninguna de las categorías extremas. El mayor acúmulo de respuestas en la categoría inferior fue en la pregunta 6 con un 27,7% de las respuestas. En la pregunta 9 se acumularon el 33,8% de las respuestas en la parte superior de la escala, representando esta la mayor frecuencia de respuestas para el “techo” de la escala. En cuanto a la puntuación total, no hubo agregación en el extremo inferior de la escala y sólo el 0,84% de las respuestas obtuvo 48 de los 48 puntos posibles en la visita de inclusión, por lo que se descartó la existencia de efectos “techo” o “suelo”.

Validez convergente o de grupos conocidos

La validez de grupos conocidos, que mide la capacidad discriminante del cuestionario, se puede valorar en la tabla 10, donde se expresan las puntuaciones medias del OHS junto con los intervalos de confianza del 95% en los diferentes terciles de la distribución de las escalas del WOMAC y el EQ-5D-5L. Las diferencias entre los tres grupos son claramente captadas por las puntuaciones del OHS con cambios medios por tercil entre 4,7 y 12,0 puntos (tabla 9).

Tabla 9. Puntuaciones medias del OHS en los pacientes clasificados según los terciles en los cuestionarios WOMAC y EQ-5D-5L.

.	Puntuación OHS para el tercil inferior de la distribución del cuestionario (IC95%)	Puntuación OHS para el tercil medio de la distribución del cuestionario (IC95%)	Puntuación OHS para el tercil superior de la distribución del cuestionario (IC95%)
WOMAC dolor	32,2 (30,7- 33,6)	22,7 (21,1- 24,2)	13,7 (12,6- 14,9)
WOMAC rigidez	29,7 (28,1- 31,2)	22,1 (20,8- 23,4)	14,4 (13,0- 15,7)
WOMAC función física	33,6 (32,3- 34,9)	21,6 (20,6- 22,6)	13,0 (11,9-14,1)
EQ-5D-5L utilidad	12,7 (11,7- 13,7)	22,4 (21,2- 23,5)	32,4 (31,1- 33,9)
EQ-5D-5L escala EVA	15,5 (14,0- 17,0)	24,0 (22,3- 25,7)	28,7 (26,7- 30,5)
OHS: Oxford Hip Score: 0 peor resultado; 48 mejor estado posible WOMAC: 0, mejor; 100 peor EQ-5D-5L Utilidades: 0, muerte ; 1, perfecto estado de salud EQ-5D EVA (0-100): 0, peor estado de salud y 100 mejor estado de salud imaginable			

En la tabla 10 se muestran los resultados para las correlaciones entre las puntuaciones del OHS con los dominios del WOMAC y las utilidades y la EVA del EQ-5D-5L. Debido a las diferentes formas de puntuación de las escalas eran esperables correlaciones negativas con el WOMAC y positivas con el EQ-5D-5L. Todas las asociaciones fueron fuertes, salvo cuando se usó la escala de rigidez del WOMAC, en cuyo caso la asociación se quedaba al borde del umbral requerido, y con la EVA del EQ-5D.

RESULTADOS

Tabla 10. Correlación (p de Pearson) entre las puntuaciones del OHS, las escalas del WOMAC y el EQ-5D-5L (índice de utilidad y escala EVA).

	OHS	WOMAC dolor	WOMAC rigidez	WOMAC función física	EQ-5D-5L Utilidad	EQ-5D-5L escala EVA
OHS	1					
WOMAC dolor	-0,807*	1				
WOMAC rigidez	-0,686*	0,698*	1			
WOMAC función física	-0,893*	0,850*	0,765*	1		
EQ-5D-5L utilidad	0,835*	-0,769*	-0,628*	-0,829*	1	
EQ-5D-5L escala EVA	0,575*	-0,563*	-0,410*	-0,560*	0,607*	1

*p<0,001
 OHS: Oxford Hip Score: : 0 peor resultado; 48 mejor estado posible
 WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.
 En el OHS y EQ-5D puntuaciones más altas indican mejor estado de salud, al contrario de lo que sucede con el WOMAC

6.2.3. Fiabilidad

El estudio de la fiabilidad incluyó la valoración de la consistencia interna mediante el alfa de Cronbach y la dimensionalidad del cuestionario y la evaluación de la repetibilidad o fiabilidad test retest mediante el coeficiente de correlación intraclass (CCI).

El alfa de Cronbach que para el OHS fue de 0,928.

Para valorar la dimensionalidad se realizó un análisis factorial exploratorio (AFE) y un análisis factorial confirmatorio.

En el AFE se encontró una estructura unidimensional con un solo factor que explicaba el 55,5% de la varianza (KMO 0,945, test de esfericidad de Bartlett χ^2 2.667, con 66 grados de libertad, p<0,001), ver tabla 11.

Tabla 11. Análisis Factorial Exploratorio: extracción de componentes principales.

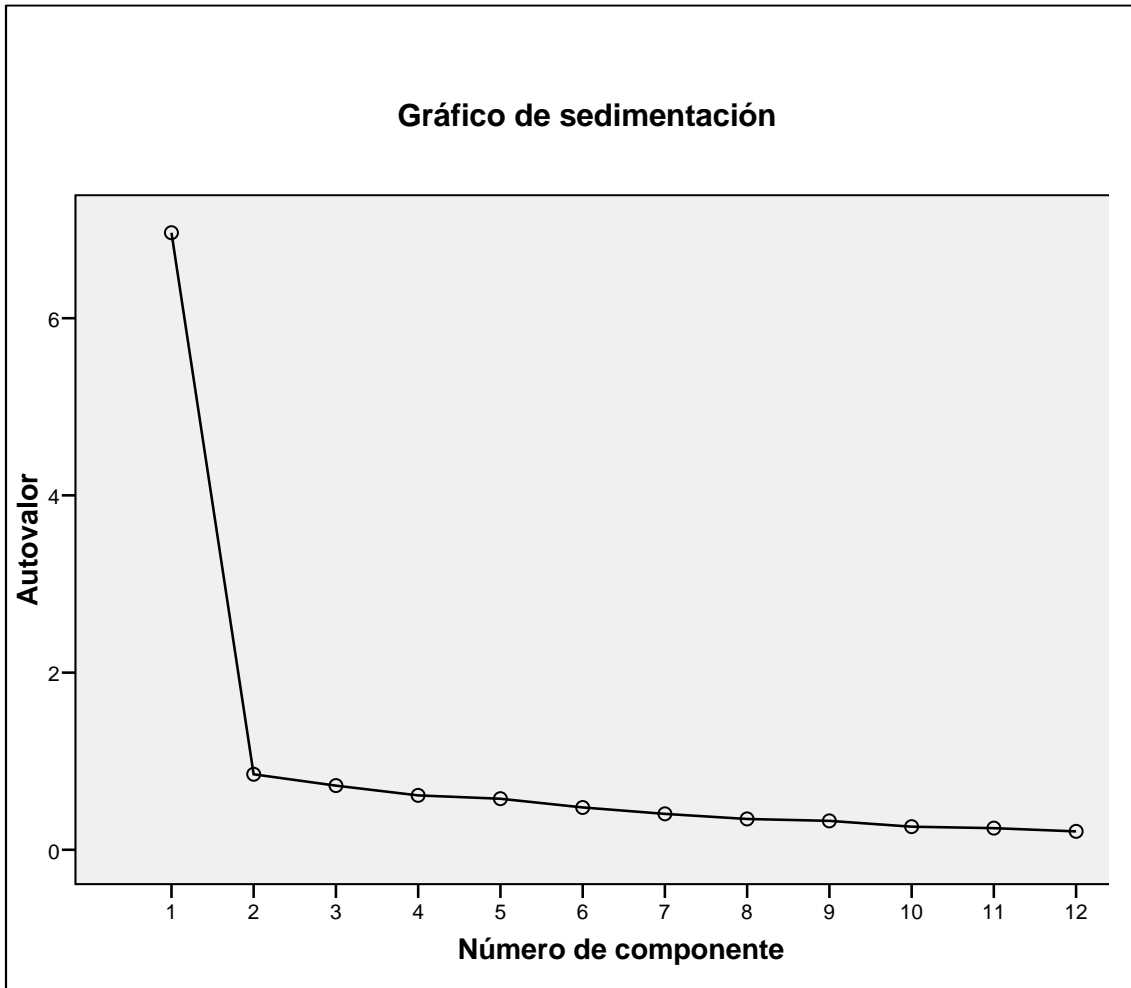
Componente	Autovalores iniciales		Sumas de las saturaciones al cuadrado de la extracción			
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	6,965	58,038	58,038	6,965	58,038	58,038
2	0,852	7,102	65,140			
3	0,724	6,036	71,176			
4	0,614	5,177	76,292			
5	0,576	4,801	81,093			
6	0,477	3,978	85,071			
7	0,405	3,378	88,449			
8	0,348	2,896	91,346			
9	0,326	2,718	94,063			
10	0,261	2,172	96,236			
11	0,244	2,034	98,270			
12	0,208	1,730	100,000			

Método de extracción: Análisis de Componentes principales

Solo hay un componente con autovalores mayores que uno. Se ve claramente en el gráfico de sedimentación.

El gráfico de sedimentación sirve para determinar el número óptimo de factores. Consiste simplemente en una representación gráfica del tamaño de los autovalores. Un autovalor indica la cantidad de varianza explicada por una componente principal (figura 2).

Figura 2. Gráfico de sedimentación del Análisis Factorial Exploratorio (AFE) del cuestionario Oxford Hip Score (OHS).



En la tabla 12 se muestran las comunalidades y los pesos factoriales para cada una de las preguntas del OHS. Mediante las comunalidades se determina la proporción de su varianza que puede ser explicada por el modelo factorial obtenido.

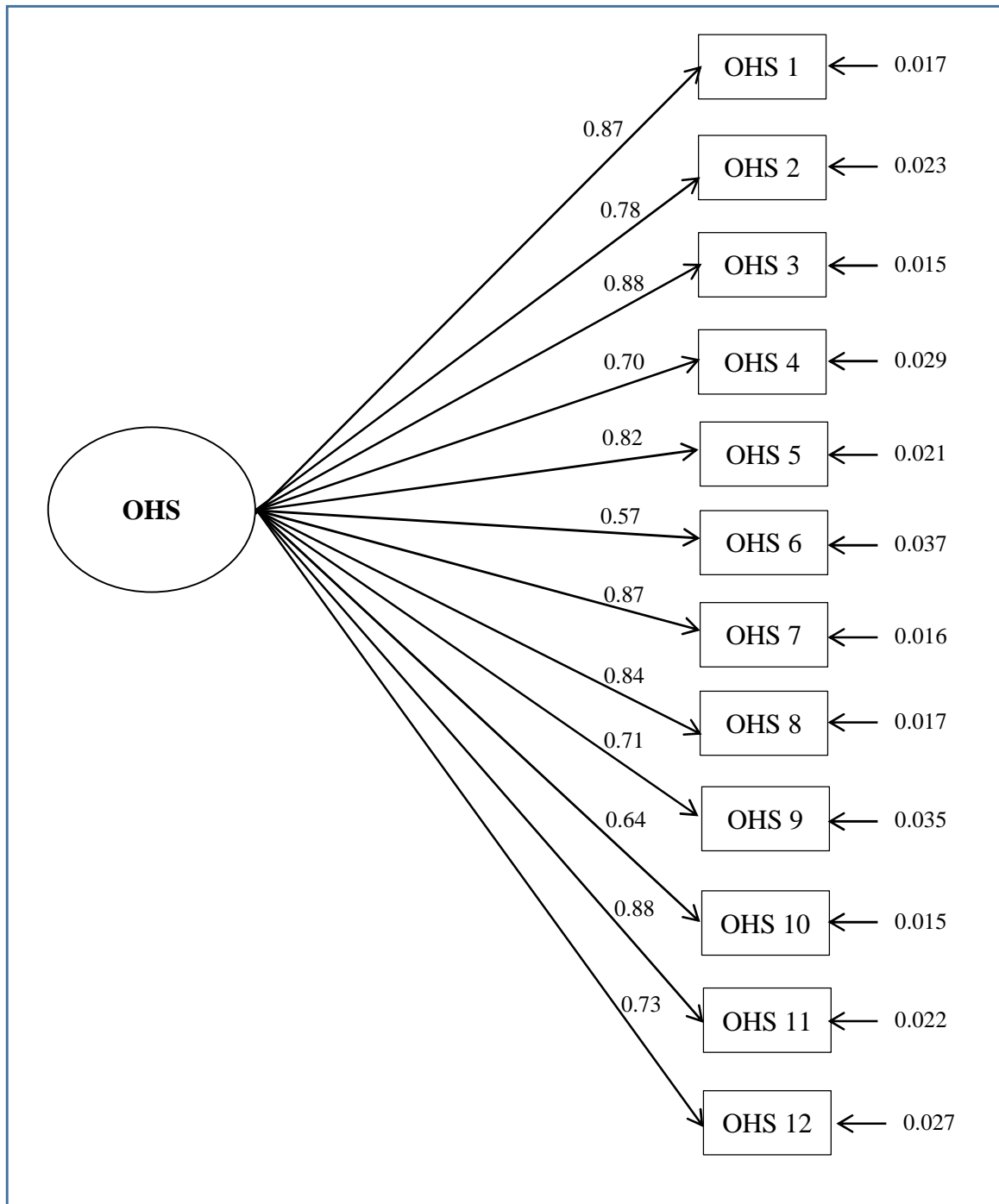
Tabla 12. Análisis factorial exploratorio para los ítems del Oxford Hip Score (OHS).

Variable	Análisis Factorial Exploratorio 1 factor	
	Peso factorial	Comunalidad
OHS 1	0,814	0,662
OHS 2	0,791	0,626
OHS 3	0,842	0,709
OHS 4	0,738	0,545
OHS 5	0,797	0,635
OHS 6	0,558	0,311
OHS 7	0,834	0,696
OHS 8	0,810	0,656
OHS 9	0,698	0,488
OHS 10	0,630	0,396
OHS 11	0,852	0,726
OHS 12	0,718	0,515
OHS: Oxford Hip Score		

Todos los pesos factoriales estuvieron por encima de 0,50 y las comunalidades fueron iguales o superiores a 0,40 salvo para las preguntas 6 y 10 (tabla 2).

En cuanto a los resultados del análisis factorial confirmatorio, los índices de ajuste fueron adecuados: (a) el RMSEA fue de 0,082, muy cercano a la cifra propuesta de 0,80; y (b) el CFI y el TLI fueron de 0,982 y 0,977, respectivamente, superando el valor de referencia de 0,95. Todas las cargas factoriales fueron estadísticamente significativas ($P < 0,001$) y oscilaron entre 0,57 y 0,88 (figura 3).

Figura 3. Análisis factorial confirmatorio para datos categóricos del cuestionario Oxford Hip Score (OHS). Se muestran los parámetros estandarizados y los errores estándar (flechas a la derecha).



Respecto a la repetibilidad del cuestionario, el coeficiente de correlación intraclase para los 124 sujetos a los que se repitió el cuestionario en los siguientes 7 a 14 días fue de 0,992 (IC95%: 0,994-0,998) y el alfa de Cronbach para las dos medidas de 0,996.

6.2.4. *Sensibilidad al cambio*

A los 6 meses, 313 sujetos pudieron ser seguidos. De ellos 65 personas habían sufrido una cirugía de remplazo articular, y 94 (30,0%, IC95%:25,0- 35,1%) manifestaron encontrarse “algo mejor” o “mucho mejor” respecto a la cadera que originó la inclusión en el estudio. Del total de sujetos seguidos, 133 (42,5%; IC95%: 37,0- 48,0%) expresaron que se sentían “algo peor” o “mucho peor”.

En la tabla 13 se muestran los resultados del cambio medio producido en las puntuaciones de los diferentes cuestionarios utilizados cuando el paciente refería que había variado su percepción de su estado de salud.

RESULTADOS

Tabla 13. Cambios en los cuestionarios OHS, EQ-5D-5L y WOMAC a los 6 meses de seguimiento en aquellos pacientes que referían cambios en su estado.

	El estado de las caderas es “mucho peor” N=53		El estado de las caderas es “algo peor” N=78		El estado de las caderas es “algo mejor” N=34		El estado de las caderas es “mucho mejor” N=60	
	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC 95%)	T.E.
OHS	-5.8 (-8.1--3.4)	0.69	-2.4 (-3.6- -1.1)	0.42	7.0 (3.7- 10.4)	0.73	18.7 (15.9-21.6)	1.71
EQ-5D-5L utilidad	-0.16 (-0.23- -0.08)	0.57	-0.06 (-0.11- -0.01)	0.09	0.19 (0.10-0.28)	0.73	0.40 (0.33- 0.47)	1.48
EQ-5D-5L escala EVA	-5.3 (-12.9- 2.2)	0.20	-1.6 (-5.6- 2.4)	0.09	9.0 (0.6- 17.3)	0.37	20.6 (14.7- 26.5)	0.91
	El dolor es “mucho peor” N=46		El dolor es “algo peor” N=77		El dolor está “algo mejor” N=42		El dolor está “mucho mejor” N=60	
	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC 95%)	T.E.
WOMAC dolor	-13.5 (-18.5- -8.4)	0.79	-4.1 (-7.2- -1.0)	0.30	14.9 (8.5-21.3)	0.73	38.3 (32.2- 44.3)	1.64
	La rigidez está “mucho peor” N=38		La rigidez está “algo peor” N=82		La rigidez está “algo mejor” N=33		La rigidez está “mucho mejor”N=56	
	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.
WOMAC rigidez	-12.2 (-20.9- -3.4)	0.50	-6.9 (-11.5- 6.2)	0.32	17.8 (8.5- 27.1)	0.68	38.6 (31.1- 46.1)	1.37
	La limitación es “mucho peor” N=48		La limitación es “algo peor” N=82		La limitación está “algo mejor” N=33		La limitación está “mucho mejor”N=47	
	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.
WOMAC función física	-8.7 (-14.5- -2.9)	0.43	-4.9 (-8.4- -1.5)	0.31	21.3 (14.0- 28.7)	1.03	43.2 (36.2- 50.2)	1.81

T.E.: Tamaño del Efecto
OHS: Oxford Hip Score : 0 peor resultado; 48 mejor estado posible
WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.
En OHS y EQ-5D se restaron puntuaciones iniciales de finales en el WOMAC al revés de forma que resultados positivos indican siempre mejoría

Cuando los sujetos referían encontrarse “algo mejor”, el tamaño del efecto para el cambio del OHS era de 0,73 y de 1,71 cuando los sujetos se encontraban “mucho mejor”. La sensibilidad para cambios negativos moderados (“algo peor”) tenía un tamaño del efecto menor, 0,42 y para cambios negativos grandes (“mucho peor”) presentaba un tamaño del efecto de 0,69. Existía un gradiente claro en las puntuaciones según el cambio percibido por el sujeto y estas son significativamente diferentes para los grupos que están “un poco peor”, “un poco mejor” y “mucho mejor”. Existe ligero solapamiento entre los que se encuentran “mucho peor” y “un poco peor”. El instrumento resultó ser más sensible que el EQ-5D-5L y funcionaba de forma parecida al WOMAC.

6.2.5. Interpretabilidad

La media del cambio de la puntuación del OHS, que se usó como estimación de la MDCl, fue de 7,0 puntos (DS 9,6), el cambio medio en sujetos con mejoría subjetiva “moderada”. El error estándar de la medición (EEM) estimado fue de 2,0, por lo que el $MDC_{95\%}$ se estimó en 5,5. Esto supone una ratio MDCl / $MDC_{95\%}$ de 1,3, esto es la herramienta es capaz de distinguir el cambio percibido por el sujeto como moderadamente importante diferenciándolo del “ruido” aleatorio.

RESULTADOS

6.3. Validación del OKS

El proceso de validación del cuestionario OKS fue similar al presentado para el OHS. Fueron incluidos un total de 397 pacientes con diagnóstico de artrosis de rodilla, 158 en Vizcaya, 158 en Madrid y 81 en Canarias. De estos un 36,8% fueron incluidos en consultas de Atención Primaria, un 57,8% en consultas de Traumatología y un 5,4% en consultas de Reumatología. El 69,8% de los pacientes eran mujeres (IC 95%: 65,3-74,3%) con una edad media para la muestra de 71,4 años (IC95%: 70,5-72,3 años).

Respecto a la localización de la OA, el 27,7 (IC95%:23,3-32,1%) tenían afectada la rodilla derecha, el 30,0 (IC95%: 25,5- 34,5%) la izquierda y el 43,3% (IC95%: 38,4-48,2%) ambas.

El 18,1% (IC95%: 14,3-21,9%) habían sufrido cirugías de remplazo total de rodilla previamente. El Índice de comorbilidad de Charlson tenía una media de 0,8 puntos (IC95%: 0,7-0,9). El IMC medio era de 29,7 (IC 95%: 29,2- 30.2).

En la tabla 14 se presentan los resultados expresados por los pacientes para los cuestionarios OKS, WOMAC y EQ-5D-5L.

Tabla 14. Puntuaciones expresadas en los cuestionarios OKS, WOMAC y EQ-5D-5L.

	N	Media puntuaciones (IC 95%)	Mediana (Recorrido Intercuartil)
Cuestionario OKS	395	22,0 (21,0- 22,9)	20 (14-27)
WOMAC dolor	396	46,2 (44,1- 48,3)	44,4 (30,0- 60,0)
WOMAC rigidez	396	45,5 (42,9- 48,0)	46,4 (25,0- 62,5)
WOMAC función física	397	52,7 (50,5- 55,0)	48,5 (25,0- 62,5)
EQ-5D-5L utilidad	393	0,54 (0,52-0,57)	0,62 (0,39- 0,74)
EQ-5D-5L escala EVA	390	57 (55,2- 59,5)	60 (45-75)
OKS: Oxford Knee Score: 0 peor resultado; 48 mejor estado posible WOMAC: 0, mejor; 100 peor EQ-5D-5L Utilidades: 0, muerte ; 1, perfecto estado de salud EQ-5D EVA (0-100): 0, peor estado de salud y 100 mejor estado de salud imaginable			

6.3.1. *Aceptabilidad*

Se obtuvo información que permitía calcular el valor resumen del cuestionario OKS en 395 casos (99,5%; IC95%: 98,8-100%). Las preguntas 7, 9 y 12 fueron contestadas en todos los casos, y las preguntas 1, 2, 3, 5, 6, 10 y 11 en todos los casos menos uno. La pregunta 8 fue dejada sin contestar en dos ocasiones y la 4 en seis cuestionarios.

6.3.2. *Validez de contenido*

“Efectos suelo” y “efecto techo”

En todas las preguntas se utilizaron todos los rangos de respuesta posible (de 0 a 4). Solo en la pregunta 7 se acumularon el 48% de las respuestas en la categoría superior. En preguntas 1 y 7 se agruparon menos del 10% de las respuestas en el extremo inferior de la escala (0 y 1) y eso no ocurrió para ninguna pregunta en su extremo superior. En cuanto a la puntuación total, no hubo agregación en el extremo inferior de la escala y sólo el 0,25% y el 0,61% de las respuestas obtuvieron 48 de los 48 puntos posibles en la visita de inclusión o en la de los seis meses, respectivamente. Por lo tanto, se descartó la presencia de efectos “suelo” o “techo”.

Validez convergente o de grupos conocidos

La validez de grupos conocidos, que mide la capacidad discriminante del cuestionario, se puede valorar en la tabla 15, donde se expresan las puntuaciones medias del OKS junto con los intervalos de confianza del 95% en los diferentes terciles de la distribución de las escalas del WOMAC y el EQ-5D-5L. Las diferencias entre los tres grupos son claramente captadas por las puntuaciones del OKS con cambios medios por tercil entre 5,6 y 11,9 puntos.

Tabla 15. Puntuaciones medias del OKS en los pacientes clasificados según los terciles en los cuestionarios WOMAC y EQ-5D-5L.

	Puntuación OKS para el tercil inferior de la distribución del cuestionario (IC95%)	Puntuación OKS para el tercil medio de la distribución del cuestionario (IC95%)	Puntuación OKS para el tercil superior de la distribución del cuestionario (IC95%)
WOMAC dolor	31,4 (30,0- 32,7)	20,5 (19,5- 21,6)	13,8 (12,7- 14,9)
WOMAC rigidez	29,7 (28,2- 31,3)	22,1 (20,8- 23,4)	14,4 (13,3- 15,4)
WOMAC función física	32,0 (30,8- 33,3)	20,1 (19,1- 21,0)	13,1 (12,1- 14,0)
EQ-5D-5L utilidad	13,6 (12,5- 14,6)	21,1 (20,0- 22,2)	31,1 (29,7- 32,4)
EQ-5D-5L escala EVA	15,7 (14,0- 17,3)	21,5 (20,3- 22,7)	27,1 (25,3- 28,9)
OKS: Oxford Knee Score: 0 peor resultado; 48 mejor estado posible WOMAC: 0, mejor; 100 peor EQ-5D-5L Utilidades: 0, muerte ; 1, perfecto estado de salud EQ-5D EVA (0-100): 0, peor estado de salud y 100 mejor estado de salud imaginable			

En la Tabla 16 se muestran los resultados para las correlaciones entre las puntuaciones del OKS con los dominios del WOMAC y las utilidades y la EVA del EQ-5D-5L. Debido a las diferentes formas de puntuación de las escalas eran esperables correlaciones negativas con el WOMAC y positivas con el EQ-5D-5L. Todas las asociaciones fueron fuertes, salvo cuando se usó la escala de rigidez del WOMAC, en cuyo caso la asociación se quedaba al borde del umbral requerido, y con la EVA del EQ-5D.

RESULTADOS

Tabla 16. Correlación (p de Pearson) entre las puntuaciones del OKS, las escalas del WOMAC y el EQ-5D-5L (índice de utilidad y escala EVA).

	OKS	WOMAC dolor	WOMAC rigidez	WOMAC función física	EQ-5D-5L Utilidad	EQ-5D-5L escala EVA
OKS	1					
WOMAC dolor	-0,745*	1				
WOMAC rigidez	-0,641*	0,643*	1			
WOMAC función física	-0,849*	0,799*	0,719*	1		
EQ-5D-5L utilidad	0,756*	-0,623*	-0,504*	-0,706*	1	
EQ-5D-5L escala EVA	0,530*	-0,472*	-0,361*	-0,516*	0,547*	1

*p<0,001
 OKS: Oxford Hip Score: : 0 peor resultado; 48 mejor estado posible
 WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.
 En el OHS y EQ-5D puntuaciones más altas indican mejor estado de salud, al contrario de lo que sucede con el WOMAC

6.3.3. Fiabilidad

La consistencia interna se evaluó mediante el alfa de Cronbach y el estudio de las dimensiones del cuestionario. Además se evaluó la fiabilidad test-retest o repetibilidad en la submuestra de pacientes con artrosis de rodilla reclutados en Madrid (n=158).

El alfa de Cronbach que para el OKS fue de 0,920.

Para valorar las dimensiones se realizó un análisis factorial exploratorio (AFE) y un análisis factorial confirmatorio.

En el AFE, de nuevo se encontró una estructura unidimensional con un solo factor que explicaba el 55,5% de la varianza (KMO 0,946, test de esfericidad de Bartlett χ^2 2.597, con 66 grados de libertad, p<0,001). Todos los pesos factoriales estuvieron por encima de 0,50 y las comunalidades fueron superiores a 0,40 salvo para las preguntas 4 y 8 (Tabla 17).

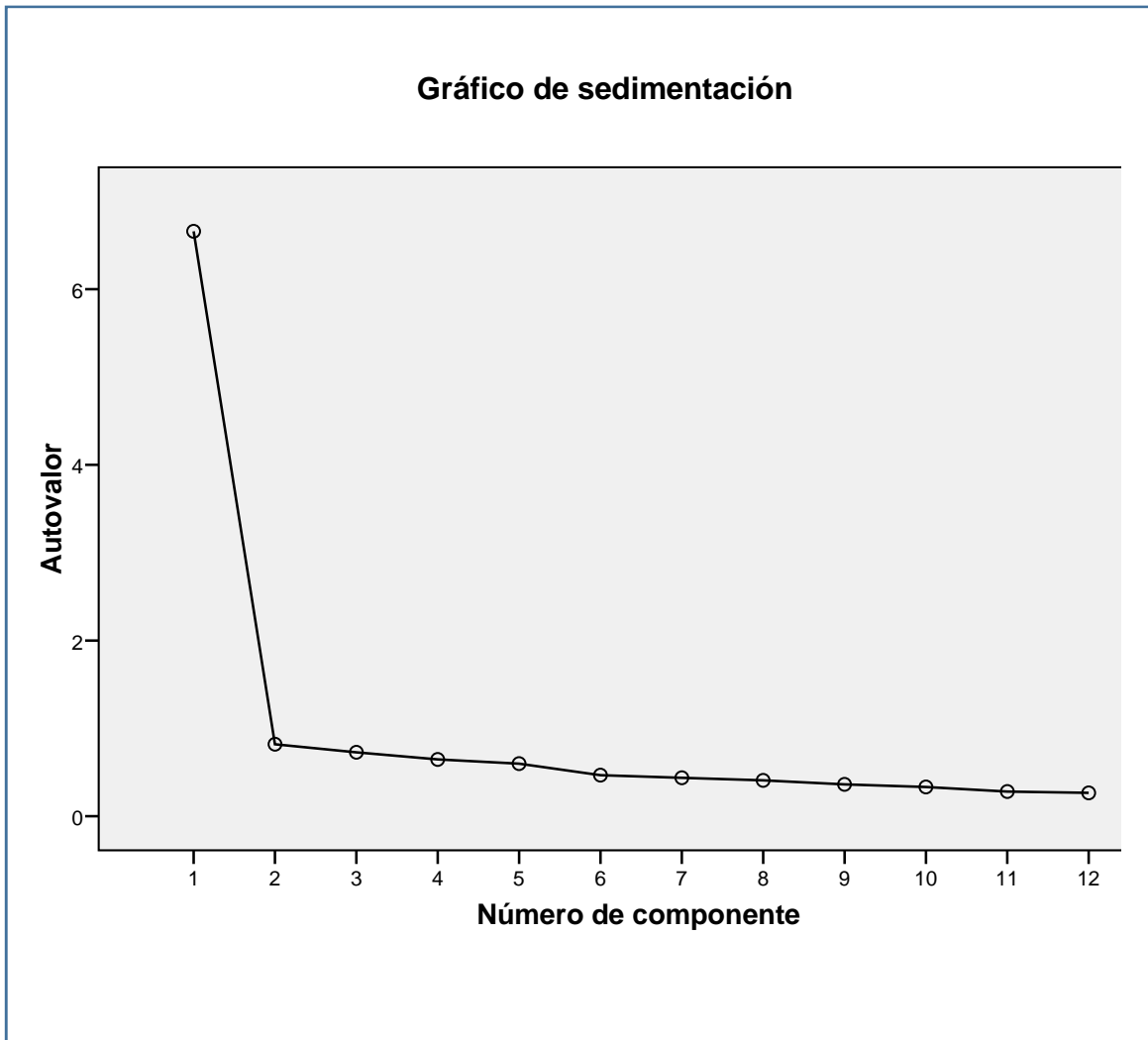
Tabla 17. Análisis Factorial Exploratorio (AFE), extracción de componentes principales autovalores para los ítems del Oxford Knee Score (OKS).-

Componente	Autovalores iniciales		Sumas de las saturaciones al cuadrado de la extracción			
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	6,659	55,492	55,492	6,659	55,492	55,492
2	0,819	6,823	62,315			
3	0,727	6,060	68,375			
4	0,646	5,385	73,760			
5	0,598	4,984	78,744			
6	0,467	3,888	82,632			
7	0,437	3,640	86,272			
8	0,408	3,398	89,670			
9	0,362	3,017	92,687			
10	0,332	2,767	95,454			
11	0,280	2,334	97,787			
12	0,266	2,213	100,000			

Método de extracción: Análisis de Componentes principales.

Solo hay un componente con autovalores mayores que uno como se ve claramente en el gráfico de sedimentación (figura 4).

Figura 4. Gráfico de sedimentación del Análisis Factorial Exploratorio (AFE) del cuestionario Oxford Knee Score (OKS).



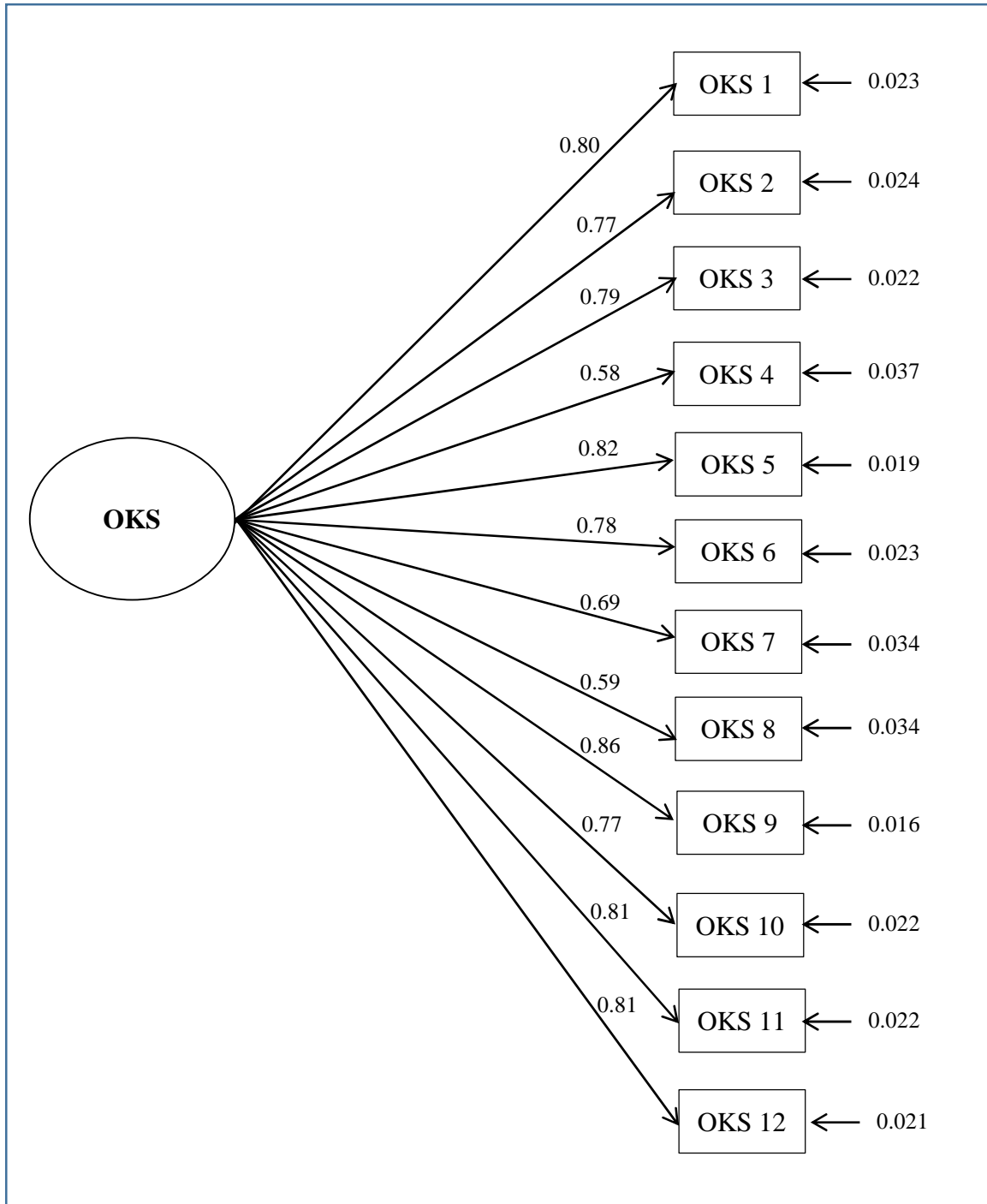
En la tabla 18 se muestran las comunalidades y los pesos factoriales para cada una de las preguntas del OHS. Como se señaló anteriormente, la comunalidad de una variable es la proporción de su varianza que puede ser explicada por el modelo factorial obtenido. Todos los pesos factoriales estuvieron por encima de 0,50 y las comunalidades fueron superiores a 0,40 salvo para las preguntas 4 y 8.

Tabla 18. Análisis factorial exploratorio (AFE) para los ítems del Oxford Knee Score (OKS).

Análisis Factorial Exploratorio 1 factor		
Variable	Peso factorial	Comunalidad
OKS 1	0,761	0,579
OKS 2	0,751	0,564
OKS 3	0,774	0,599
OKS 4	0,572	0,327
OKS 5	0,796	0,634
OKS 6	0,773	0,597
OKS 7	0,670	0,449
OKS 8	0,605	0,365
OKS 9	0,839	0,705
OKS 10	0,767	0,589
OKS 11	0,788	0,620
OKS 12	0,794	0,630
OKS: Oxford Knee Score		

En cuanto a los resultados del análisis factorial confirmatorio, los índices de ajuste fueron excelentes: (a) el RMSEA fue de 0,076, inferior a 0,08; y (b) el CFI y el TLI fueron de 0,981 y 0,977, respectivamente, superando el valor de referencia de 0,95. Las cargas factoriales fueron todas estadísticamente significativas ($P < 0,001$) y oscilaron entre 0,58 y 0,86 (Figura 5).

Figura 5. Análisis factorial confirmatorio para datos categóricos del cuestionario Oxford Knee Score (OKS). Se muestran los parámetros estandarizados y los errores estándar (flechas a la derecha).



Para evaluar la repetibilidad se utilizó el CCI para los 158 sujetos a los que se repitió el cuestionario en los siguientes 7 a 14 días, que fue de 0,993 (IC95%: 0,990-0,995) y el alfa de Cronbach para las dos medidas de 0,997.

6.3.4. Sensibilidad al cambio

A los 6 meses, 331 sujetos pudieron ser seguidos. De ellos 42 personas habían sufrido una cirugía de remplazo articular, y 101 (30,5%, IC95%:25,6- 35,5%) manifestaron encontrarse “algo mejor” o “mucho mejor” respecto a la rodilla que originó la inclusión en el estudio. Del total de sujetos seguidos, 143 (43,2%; IC95%: 37,9- 48,5%) expresaron que se sentían “algo peor” o “mucho peor”.

En la Tabla 19 se muestran los resultados del cambio medio producido en las puntuaciones de los diferentes cuestionarios utilizados cuando el paciente refería que había variado su percepción de su estado de salud. Cuando los sujetos referían encontrarse “algo mejor”, el tamaño del efecto para el cambio del OKS era de 0,69 y de 1,60 cuando los sujetos se encontraban “mucho mejor”. La sensibilidad para cambios negativos moderados (“algo peor”) tenía un tamaño del efecto menor, 0,24 y para cambios negativos grandes (“mucho peor”), presentaba un tamaño del efecto de 0,57. Existía un gradiente claro en las puntuaciones según el cambio percibido por el sujeto y estas son significativamente diferentes para los grupos que están “un poco peor”, “un poco mejor” y “mucho mejor”. Existe ligero solapamiento entre los que se encuentran “mucho peor” y “un poco peor”. El instrumento resultó ser más sensible que el EQ-5D-5L y funcionaba de forma parecida al WOMAC en las escalas de dolor y limitación, mientras que la escala de rigidez era menos sensible al cambio.

RESULTADOS

Tabla 19. Cambios en los cuestionarios OKS, WOMAC y EQ-5D-5L a los 6 meses de seguimiento en aquellos pacientes que referían cambios en su estado.

	El estado de las rodillas es “mucho peor” N=59		El estado de las rodillas es “algo peor” N=84		El estado de las rodillas es “algo mejor” N=56		El estado de las rodillas es “mucho mejor” N=45	
	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC 95%)	T.E.
OKS	-4.5 (-6.6--2.4)	0.57	-1.6 (-2.9- -0.2)	0.24	6.1 (3.7- 8.5)	0.69	17.4 (14.1-20.7)	1.60
EQ-5D-5L utilidad	-0.17 (-0.25- -0.08)	0.55	-0.01 (-0.06- 0.03)	0.07	0.14 (0.07-0.20)	0.43	0.30 (0.21- 0.39)	0.96
EQ-5D-5L escala EVA	-4.6 (-12.8-3.7)	0.15	-0.74 (-4.9- 3.4)	0.04	7.9 (3.0- 12.8)	0.58	13.2 (9.0-17.4)	0.94
	El dolor es “mucho peor” N=51		El dolor es “algo peor” N=82		El dolor está “algo mejor” N=61		El dolor está “mucho mejor” N=47	
	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC 95%)	T.E.
WOMAC dolor	-9.9 (-14.9- -4.8)	0.55	-2.0 (-5.2- 1.3)	0.13	17.7 (13.1-22.3)	0.98	28.8 (21.5-36.1)	1.15
	La rigidez está “mucho peor” N=44		La rigidez está “algo peor” N=85		La rigidez está “algo mejor” N=52		La rigidez está “mucho mejor” N=42	
	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.
WOMAC rigidez	-13.1 (-20.9- -5.3)	0.51	-0.7 (-5.9- 4.5)	0.03	12.3 (5.1- 19.4)	0.48	30.1 (21.9- 38.2)	1.15
	La limitación es “mucho peor” N=47		La limitación es “algo peor” N=90		La limitación está “algo mejor” N=47		La limitación está “mucho mejor” N=31	
	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.	Cambio medio (IC95%)	T.E.	Cambio medio (IC 95%)	T.E.
WOMAC función física	-10.5 (-15.2- -5.7)	0.65	-2.4 (-5.5- 0.7)	0.16	18.4 (12.3- 24.4)	0.89	38.6 (29.7- 47.6)	1.58
T.E.: Tamaño del Efecto								
OKS: Oxford Knee Score : 0 peor resultado; 48 mejor estado posible								
WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.								
En OHS y EQ-5D se restaron puntuaciones iniciales de finales en el WOMAC al revés de forma que resultados positivos indican siempre mejoría								

6.3.5. Interpretabilidad

La media del cambio de la puntuación del OKS en sujetos con mejoría subjetiva “moderada”, que se usó como estimación de la MDCl, fue de 6,1 puntos (DS 8,9). El EEM estimado fue de 1,5, por lo que el $MDC_{95\%}$ se estimó en 4,38. Esto supone una ratio MDCl / $MDC_{95\%}$ de 1,4.

Esto indica que la herramienta permite identificar el cambio percibido por el sujeto como moderadamente importante diferenciándolo del “ruido” aleatorio.

RESULTADOS

7. Discusión

DISCUSIÓN

En este trabajo se ha puesto de manifiesto la importante repercusión de la artrosis de miembro inferior sobre la calidad de vida y se han evaluado las propiedades psicométricas de dos nuevas herramientas adaptadas a nuestro medio cultural para valorar dicha CVRS en pacientes con artrosis de cadera y rodilla.

Las características evaluadas posicionan al OHS y al OKS, en sus versiones adaptadas al castellano como herramientas adecuadas tanto para discriminar la situación de pacientes distintos, como para valorar la evolución de un paciente o grupo de pacientes.

El grupo de pacientes incluidos en el estudio parece adecuado para conseguir los objetivos perseguidos. La muestra incluida fue lo suficientemente amplia como para haber recogido correctamente el espectro de la enfermedad en todos los niveles asistenciales. Además también parece adecuada su representatividad respecto a la población consultante de todo el país.

Si bien es cierto que no tenemos datos de la población que no utiliza el Sistema Nacional de Salud (SNS), en España, el 99,0 % de los ciudadanos reportaron tener cobertura por dicho sistema (361). En el año anterior a esta encuesta se produjeron 2.086,3 consultas especializadas ambulatorias por cada 1000 habitantes y 239,2 millones de visitas a médicos de atención primaria con una media de 5,2 visitas por persona (362), lo que permite pensar que la población consultante en el SNS es representativa de la población general.

La submuestra de pacientes mayores de 50 años, no mostró diferencias relevantes respecto a la muestra total y permitía tener un conjunto de pacientes más homogéneo y comparable con otros estudios sobre la prevalencia de artrosis de cadera y rodilla en la población española (101) empleando los criterios clínicos más utilizados y mejor validados (94,331).

En la población de estudio con OAMI nos encontramos que predominan las mujeres, representando hasta el 62% de la muestra, pero con una distribución desigual para la cadera y la rodilla. En el caso de la artrosis de rodilla el sexo femenino representa casi el 70% de la muestra total, mientras que para la artrosis de cadera las proporciones se igualan, representado el 53% las mujeres. Estos datos se muestran acordes con los estudios descriptivos de los factores de riesgo de la artrosis, donde el sexo femenino se asocia con una mayor prevalencia de la artrosis en particular de la rodilla (17), documentándose que el riesgo de padecerla es casi un 70% superior en las mujeres (120), mientras que no se observan diferencias significativas por sexo para la coxartrosis (12).

DISCUSIÓN

En cuanto a la edad, encontramos un aumento de la prevalencia de la artrosis que se incrementa con la edad desde los 50 años hasta los 80 años, con un pico de prevalencia en la séptima década de la vida. Los estudios determinan que la edad se muestra como el principal factor de riesgo en la artrosis en sus diferentes localizaciones, de tal forma que la prevalencia y la incidencia de la artrosis radiológica aumentan de forma considerable con la edad (15). Aunque la tendencia no es lineal, entre los 50-75 años se observa un fuerte aumento de la incidencia, que es más limitada después de los 75 años (120). En nuestra muestra observamos este aumento en la prevalencia desde los 50-54 años, población que representa el 4,7% de la muestra, hasta el grupo de los 75-79 años, que representan el 21% de la muestra.

Los pacientes con OAMI de nuestra muestra pertenecen a los grupos sociales más desfavorecidos que se corresponden con los grupos V y VI de la clasificación de clase social ocupacional neoweberiana propuesta por Domingo-Salvany y cols (339). En el caso de los pacientes con artrosis de cadera, el 22,4% desarrollan actividades en el sector primario y el 14,7% desarrollan actividades no cualificadas, lo que representa el 37,1% de la muestras. Cifras muy similares para la artrosis de rodilla donde los pacientes que desarrollan actividades del sector primario representan el 22,7%, y el 16,9% desarrollan actividades no cualificadas, ambos grupos suman casi el 40% de la muestra. Encontramos numerosos estudios en la bibliografía sobre la influencia de las tareas profesionales con la aparición de artrosis. Los trabajos físicamente exigentes se han asociado fuertemente con un mayor riesgo de artrosis de cadera y rodilla (80,363). Así como adoptar determinadas posiciones biomecánicamente estresantes para las articulaciones, como arrodillarse, trabajar de cuclillas o permanecer mucho tiempo de pie (38,82).

Los pacientes estudiados también presentaban un elevado índice de masa corporal, que sitúa a más del 80% en una situación de sobrepeso u obesidad. El IMC medio de los pacientes con artrosis de cadera es superior a 28, y mayor de 29 para los pacientes con artrosis de rodilla. Es conocido que el sobrepeso u obesidad en sus diferentes grados destaca como uno de los factores asociados al estilo de vida que aumentan la probabilidad para desarrollar OAMI; siendo la asociación más evidente para la artrosis de rodilla (20,120), y no lo es de forma significativa en el caso de la artrosis de cadera (21). En nuestra muestra podemos observar como los pacientes con OAMI presentan obesidad en un porcentaje 10 puntos superior, del 27,7% al 37% comparado con la población general que tomamos como referencia, y que ha sido extraída de una muestra representativa de la población española de 50 años.

La afectación articular múltiple es frecuente en los sujetos de nuestra muestra. Dentro del grupo de pacientes con artrosis de cadera, encontramos que el 28,2% de los casos presentan afectación

de en ambas caderas, y el 31,8% afectación de la rodilla. En el grupo de pacientes con artrosis de rodilla, la artrosis en otras articulaciones está presente hasta en el 70% de los casos, de los cuales el 13% corresponden a afectación en forma de coxartrosis. Se ha demostrado el hecho de que tener artrosis de cadera multiplica el riesgo de desarrollar gonartrosis, del mismo modo, que la presencia de artrosis de rodilla multiplica el riesgo de coxartrosis, por tanto, la artrosis en sí misma se comporta como un factor de riesgo para desarrollar artrosis en otra articulación (120,328). Estos datos apoyan la idoneidad de la recogida de información clínica realizada y la potencial representatividad de la población incluida.

Una vez discutidos estos aspectos, se valoraran los resultados recogidos en la sección correspondiente.

7.1. Impacto de la artrosis de miembro inferior en la CVRS

Podemos considerar el estudio que se presenta, como el primero en nuestro medio que cuantifica la pérdida de utilidad atribuida a la OAMI, ajustada por otras variables que se conoce se asocian con cambios en la percepción de la CVRS como la edad, el número de enfermedades crónicas, el sexo y otros condicionantes sociales como el grupo social o la situación de potencial soledad. Además utiliza para medir este impacto el algoritmo propuesto para informar los resultados de las evaluaciones económicas de tecnologías (190), con la última versión del EQ-5D-5L, una herramienta con ventajas sobre versiones anteriores y cuya versión en castellano ha demostrado excelentes propiedades psicométricas en pacientes con OAMI (173). La OAMI se asocia a una enorme disminución en la utilidad atribuida al estado de salud, una vez ajustada esta asociación, por otras variables que pueden confundirla. Vivir con OAMI implica una pérdida de utilidad de alrededor de un 0,30 de media, esto es, haciendo una aproximación muy general, sería igual de preferible vivir diez años con OAMI que siete en estado de salud perfecta. Por otra parte también se muestra, como se podía intuir, que la población atendida en el nivel especializado tiene peor percepción de su CVRS que la atendida en atención primaria, de media 0,11 puntos inferior en la atribución de utilidades, lo que equivale a una pérdida aproximada de un año de vida en buena salud por cada diez años vividos.

La afectación articular degenerativa se había manifestado como una de las afectaciones crónicas que más afectan la calidad de vida en la población general en Reino Unido, solo por detrás del dolor y la ansiedad/depresión (364). Datos similares se habían comunicado para la población canadiense (107). La pérdida de utilidad atribuida en estos estudios a la artrosis en general oscilaba alrededor de 0,10 puntos cuando se ajustaba por otras variables que podían afectar a la CVRS percibida, una pérdida unas tres veces superior a la producida por otras condiciones crónicas como la diabetes o el asma (107). En España, el estudio de la artrosis en su conjunto la sitúa como una condición crónica que disminuye la utilidad en la misma medida que los problemas cardiacos o la diabetes mellitus, con una disminución de las utilidades inferiores a 0,10 puntos respecto a población general mayor de 65 años cuando se utiliza el EQ-5D-3L para evaluar la CVRS (130). Cuando nos referimos en concreto a la OAMI, se han comunicado en la literatura impactos más importantes sobre las utilidades atribuidas a estos estados de salud, similares a los encontrados en el presente estudio. En una población italiana se veía una pérdida de utilidad en pacientes con

OAMI de aproximadamente 0,28 puntos de media respecto a población general (365). En ese caso se evaluó la utilidad con la versión de 3 niveles del EQ-5D y solo se produjo un ajuste por edad y sexo. Según estos autores, el impacto de la OAMI sobre las utilidades es similar al producido por la osteoporosis con fractura vertebral o la espondilitis anquilopoyética, y muy superior al que producían otras enfermedades crónicas que afectan al sistema locomotor como el Síndrome de Sjogren o la esclerosis sistémica (365). Estas cifras superan con mucho los umbrales de sensibilidad del EQ-5D-5L, obtenidos en nuestro medio (0,07 para sujetos con OAMI no intervenidos) (173), o en otros países (0,09 para sujetos con artritis) (107).

También se ha podido estudiar la asociación entre el nivel en que es captado el paciente, y la situación de salud expresada. Los pacientes atendidos en el ámbito hospitalario expresan un nivel de utilidad media 0,11 puntos inferior. En nuestro país se había valorado la CVRS en pacientes con OAMI en consultas de especialistas en traumatología (366), y en el medio comunitario (367), con resultados no comparables por el uso de diferentes herramientas de medida. Esta es la primera vez que se presenta una evidencia sobre el diferente grado de afectación de la enfermedad, cuando el paciente acude a diferentes niveles asistenciales, y el resultado tiene coherencia con la organización del sistema nacional de salud español que otorga al primer nivel asistencial la función de filtro respecto a la atención de diferentes enfermedades crónicas y de la artrosis en particular (362).

Puede que este gran impacto de la OAMI sobre la CVRS se deba a que afecta a varias dimensiones de este constructo, incorporadas en la herramienta EQ-5D, como el dolor y la pérdida de función (368). Pero los datos presentados muestran, en este caso, una afectación de un componente psicológico, capturado en la dimensión ansiedad/depresión del EQ-5D-5L. Este impacto de la OAMI en la salud mental ya había sido descrito en la literatura (132). Además la dimensión salud mental como componente de la CVRS percibida se ha mostrado como factor asociado al devenir del recambio articular en la OAMI (369), por lo que debería ser evaluado con especial atención.

El resto de características sociodemográficas y clínicas, que ajustan la asociación entre la OAMI y las utilidades de los estados de salud, se comportan de la manera que se conocía previamente. Las personas con mayor edad, las de grupos socioeconómicos más desfavorecidos, las que viven solas, las mujeres, las personas obesas y las que presentan más condiciones crónicas, suelen tener peor percepción del estado de salud, como se ha descrito en anteriores trabajos en nuestro país (160,161). Ya se había descrito que la asociación entre la artrosis y la CVRS reportada, se veía confundida por otras variables como los factores sociodemográficos, la enfermedad crónica o la

DISCUSIÓN

existencia de obesidad (364,367), por lo que los resultados presentados pueden considerarse una aproximación realista al estudio de asociación entre OAMI y CVRS.

No había diferencias significativas entre la CVRS o el nivel de gravedad, medido con los instrumentos específicos entre los pacientes con artrosis de cadera o rodilla. No se encontraron diferencias en la puntuación media del WOMAC para la cadera (50,85) y para la rodilla (49,95). Tampoco hubo diferencias, en las comparaciones brutas en las puntuaciones medias en la escala de Oxford para la cadera (22,72), frente a la rodilla (21,97). Exclusivamente en la EVA del EQ-5D-5L se apreciaba que los pacientes con artrosis de cadera podían tener una peor percepción que los diagnosticados de artrosis de rodilla aunque la diferencia quedaba al borde de la significación estadística, con puntuaciones en EQ-5D-5L EVA de 54,92 para la cadera y 57,47 en el caso de la artrosis de rodilla.

Sin embargo, esta aparente similitud en la percepción de la CVRS en pacientes con artrosis de cadera y rodilla, podía estar escondiendo una diferente distribución de potenciales factores confusores en ambos grupos. Así se demuestra al valorar los resultados de los modelos presentados. Cuando solo se estudia el grupo con diagnóstico de artrosis (modelos 1', 2' y 3') se comprueba que la artrosis de rodilla supone un menor impacto sobre la CVRS que la artrosis de cadera con puntuaciones medias de utilidad de 0,055 superiores a los pacientes con artrosis de cadera ajustado por edad, IMC, lugar de atención y existencia de artrosis bilateral, puntuación muy cercana a los 0,07 puntos que constituyen la MDCI del EQ5D en estos pacientes (173), lo que hace inferir que estas diferencias no pueden deberse solo a ruido aleatorio. Esta diferencia ya se había reportado en otros estudios, que muestran que la rigidez es mayor y la limitación funcional peor en los pacientes con artrosis de cadera o de ambas localizaciones, que en los pacientes con artrosis de rodilla. Además clínicos y pacientes perciben que la artrosis de cadera supone más dolor corporal, peor percepción de salud general y de función social, que la artrosis de rodilla (366).

El lugar de atención es importante a la hora de valorar a los pacientes, como se desprendía del estudio comparativo con la población sin artrosis. Se ve incluso una asociación más fuerte (el coeficiente pasa de 0,112 a 0,133 (con signo negativo) en los modelos definitivos. Como se señaló, este fenómeno es esperable en un entorno de Sistema Público de Salud en el que el nivel asistencial de Atención Primaria hace de "filtro" de las derivaciones a otros niveles asistenciales, manteniendo la atención de las patologías más abordables en el ámbito comunitario (362).

Ser varón se asocia a una mejor percepción de la CVRS medida a través de las utilidades (asociación al borde de la significación estadística) y tanto el bajo peso, como la obesidad se asocian a peor percepción de CVRS en pacientes con artrosis de miembro inferior.

Es curioso que, al analizar la variabilidad de la percepción de la CVRS en pacientes con artrosis de miembro inferior ya diagnosticada, no aparezcan como explicativas variables tradicionalmente asociadas a la percepción de la CVRS como la edad (160,370) o el grupo social (160,371).

Respecto a la edad, señalar que se ha descrito un fenómeno de “acostumbramiento” a situaciones de salud que se asocian al envejecimiento (372), que hacen que se valore de forma diferente dicha situación en diferentes etapas vitales.

La pérdida de influencia del grupo social sobre la CVRS en pacientes que ya tienen instaurada la enfermedad artrósica no es contradictorio con el hecho de que la enfermedad tenga una distribución claramente diferenciada en grupos sociales. De hecho la pérdida de significación de esta asociación debería tomarse con precaución una vez establecida su existencia cuando se hace una comparación con la población general con tamaños de muestra más grandes en la primera parte del estudio analizado previamente. En la literatura se han reportado asociaciones entre el nivel educativo (un “proxy” del grupo social) y la enfermedad artrósica de miembro inferior y se ha demostrado que vivir en una zona con privación económica predice la aparición de la restricción laboral en trabajadores mayores con OAMI (373).

7.1.1. Limitaciones

El diseño de este trabajo presenta algunas limitaciones. Los estudios transversales son problemáticos a la hora de establecer asociaciones causales aunque se ha valorado el efecto de los principales factores de confusión recogidos en la literatura. Por otra parte la muestra utilizada de pacientes con OAMI no puede entenderse como representativa de la población afectada de esta condición, sino que fue recogida con criterios de oportunidad. Sí podemos afirmar que la inclusión se realizó en sitios diversos de la geografía y que el perfil de los pacientes con OAMI estudiados, más mayores, con peor nivel socioeconómico, con mayor prevalencia de obesidad y predominantemente mujeres, coincide con el comunicado en otros estudios españoles (374) y europeos (367). El hecho de que los pacientes podían tener más gravedad que los que se encuentran en la población general, por la sobrerrepresentación del ámbito hospitalario, fue tenido en cuenta en el análisis. El planteamiento de la existencia de afecciones crónicas podría

DISCUSIÓN

considerarse una limitación del estudio, ya que fueron comunicadas por el sujeto y esta información podía estar sometida a sesgos de memoria. Sin embargo, a pesar de los problemas relacionados con los datos de la población general, la ENSE constituye la información más representativa y de mayor calidad disponible en ese momento. Por otra parte, el EQ-5D-5L es un instrumento genérico diseñado para medir dimensiones de la salud relevantes para todos los estados de salud, incluidos los individuos sanos, y no la percepción de los pacientes sobre aspectos concretos de la salud específicamente afectados por la OAMI. Pero tiene la ventaja de permitirnos atribuir las preferencias de los pacientes al estado de salud descrito en los pacientes con OAMI y compararlas entre diferentes poblaciones (375).

7.1.2. Implicaciones

La importancia de los resultados presentados radica en el interés que pueda tener para el decisor en políticas de salud, el hecho de tener cuantificado, en una medida reproducible y comparable, el impacto que tiene la OAMI sobre la CVRS del sujeto. Estos resultados permiten cuantificar, conociendo la prevalencia de la enfermedad, la carga de enfermedad que supone, desde la perspectiva del paciente. Además, existen intervenciones, como las cirugías de remplazo articular, que suponen una inversión considerable de recursos, pero que pueden aportar unos resultados muy valorables desde la perspectiva del paciente y del conjunto de la sociedad. Se señala continuamente que, para valorar el resultado de dichas intervenciones, es necesario tener medidas de resultados informadas por el paciente (PROM), como la calidad de vida, que permitan determinar cuán prioritaria pueda ser determinada actuación (233).

7.2. Validación del Oxford Hip Score (OHS)

El cuestionario OHS en su versión en castellano es una herramienta válida para medir la CVRS en pacientes con artrosis de cadera, es fiable y sensible al cambio. Además es una herramienta aceptable por la población a la que se dirige, pues la tasa de respuestas es extraordinariamente alta.

El estudio de la validez del OHS se ha abordado desde varias perspectivas. La validez “aparente” no se evaluó al tratarse de una adaptación.

La validez de grupos conocidos (validez discriminativa) parece correcta expresando valores diferentes los sujetos de grupos que son clasificados en diferentes niveles de CVRS con cuestionarios específicos o generales. Además no parece tener efectos techo y suelo relevantes que comprometan dicha capacidad discriminativa, como ya se había señalado para la versión original (376), y con otras adaptaciones (254,255).

El estudio de la validez convergente arrojó correlaciones con las escalas específicas del WOMAC y genéricas del EQ-5D-5L más fuertes que las encontradas en la versión original con herramientas genéricas de CVRS (376), y similares o ligeramente superiores a las encontrada para adaptaciones del OHS a otros idiomas como el alemán (255), el neerlandés (253), francés (254), italiano (256), danés (257), turco (258), coreano (259), japonés (260), rumano (261), y recientemente al árabe (262).

La validez convergente en las adaptaciones del OHS a diferentes entornos culturales muestra correlaciones más o menos fuertes. La correlación con el WOMAC es fuerte en este caso y lo mismo ocurre con la adaptación al turco (258). En ambos trabajos la correlación es fuerte con valores de r de Pearson o la rho de Sperman por encima de 0,7 (323). La adaptación al coreano explora las correlaciones del OHS con el HOOS y estas también son fuertes (259). La adaptación del OHS al rumano encuentra correlaciones fuertes entre éste y el HHS y moderadas con la EVA del EQ-5D (261), como ocurre en nuestro trabajo.

En el análisis de la fiabilidad, la consistencia interna, el alfa de Cronbach fue mejor que el encontrado para la escala original en el periodo de inclusión (0,93 vs 0,84) (249). Aunque un valor muy alto del alfa de Cronbach podría indicar que los ítems son redundantes, es poco probable que sea así, ya que se descartó mediante el análisis factorial. Este coeficiente es útil para estimar

DISCUSIÓN

la fiabilidad, especialmente en el caso de una prueba unidimensional. Si una prueba muestra un valor elevado del alfa de Cronbach, puede concluirse que su varianza es atribuible en gran medida a factores generales y de grupo. Cuando se ha demostrado la existencia de un único factor, entonces el alfa de Cronbach puede utilizarse para concluir que el conjunto de ítems es unidimensional (355).

La fiabilidad test-retest, medida mediante el coeficiente de correlación intraclase, fue elevada, y los valores obtenidos superiores a 0,90, en una muestra de más de 150 pacientes permiten calificar a la herramienta como fiable (323). Estos valores son similares a los encontrados en otras herramientas validadas al castellano para medir la CVRS en artrosis de miembro inferior, como la escala funcional de las extremidades (LEFS) (219) o el Hip Outcome Score (HOS) (236), y superiores a los encontrados para la validación al castellano del cuestionario OAKHQOL (Osteoarthritis Knee and Hip Quality of Life), cuyo análisis de repetibilidad arrojaba valores para el CCI entre 0,75 y 0,81 (225). En otros casos, como en la validación del cuestionario de resultados de cadera y rodilla (Hip and Knee Outcomes Questionnaire), no se proporcionó la fiabilidad test-retest (223).

El estudio de la fiabilidad de esta versión adaptada del OHS arrojó valores para la consistencia interna y la fiabilidad similares a la otra herramienta validada en castellano, el Hip Outcomes Score, también diseñado para valorar el cambio en la percepción de la CVRS tras intervenciones quirúrgicas, en pacientes con artrosis de cadera (236).

En cuanto a su dimensionalidad, la estructura factorial del OHS ha sido discutida. Algunos autores propusieron su descomposición en dos dimensiones, dolor y funcionalidad (377). Cuando se ha intentado comprobar si era mejor una estructura unidimensional o bidimensional, los resultados indicaban que hay motivos para aceptar la estructura uni o bidimensional, aunque en este último caso, varios ítems saturaban en ambos factores (378). Por estos motivos y a la vista de nuestro análisis factorial exploratorio, decidimos testar la estructura unidimensional, que parece una solución correcta para esta versión adaptada, ya que en el AFC se obtuvieron valores del RMSEA cercanos al límite de la aceptabilidad, con índices TLI y CFI óptimos (353).

Los ítems de OHS se generaron mediante la realización de entrevistas cualitativas con pacientes antes y después de someterse a una cirugía de reemplazo de cadera, lo que sugirió que el dolor y la discapacidad funcional estaban generalmente inextricablemente vinculados.

El análisis factorial exploratorio demostró que el OHS de dos factores tenía la mayoría de los ítems cargando de manera sobresaliente en cualquiera de los dos factores. Esto es de esperar, ya que en ciertos contextos (como la artrosis avanzada o en el momento de la artroplastia), se ha demostrado que el dolor y la función tienen una superposición considerable. Por tanto, además del uso actual del OHS como una única escala de resumen de CVRS, en otros entornos se ha demostrado que se puede extraer información separada sobre el dolor y la función de forma significativa en forma de subescalas (378).

La capacidad discriminativa del cuestionario, la que hace referencia a su potencial para distinguir entre sujetos con situaciones diferentes, es adecuada, pero la herramienta también puede ser utilizada para estudiar el cambio de percepción respecto a su situación de un mismo sujeto, esto es, tiene una adecuada capacidad evaluativa (147). El instrumento fue diseñado para ese fin y la versión adaptada al castellano ha demostrado su potencialidad, no solo en pacientes que son intervenidos de la cadera sino en la evolución de una cohorte de pacientes con artrosis de cadera en el corto plazo.

El tamaño del efecto para cambios positivos “moderados” estuvo un poco por debajo del umbral de 0,8, y en valores similares al WOMAC. El tamaño del efecto detectado para los cambios en el trabajo de validación original tras cirugía de 1,93 (379), solo puede compararse con las mejoras muy importantes (tamaño del efecto 1,71), pues en este caso se testó en un conjunto de pacientes con abordajes conservadores o quirúrgicos. El OHS había demostrado ser superior a otros cuestionarios, como el WOMAC y a los genéricos como el EQ-5D, cuando se valoran cambios grandes, como ocurre tras la cirugía de remplazo de cadera (376). En este trabajo el OHS mostró una capacidad similar al WOMAC para detectar cambios “moderados”, y ligeramente superior para cambios “grandes”.

Atendiendo al cambio reportado por los sujetos que notan mejoría “moderada” la MDCl fue de 7,0 puntos. La MDCl estimada para la versión original del OHS y calculada a partir de los valores reportados por pacientes sometidos a cirugía de remplazo articular fue de 7,5 puntos (380). Si se utiliza el criterio de estimar la MDCl a través de valorar la mitad de una desviación estándar de la distribución de las puntuaciones en los sujetos que cambian (381), podríamos situar la MDCl alrededor de los 4,8 puntos. Este criterio había situado la MDCl en valores similares, entre 3 y 5 puntos, en la validación de la escala original (250).

DISCUSIÓN

La capacidad evaluativa, como en otros cuestionarios, es superior para detectar cambios positivos que negativos (382), aunque incluso en los cambios negativos tiene una capacidad similar o superior al WOMAC y, por supuesto, al cuestionario genérico EQ-5D-5L.

El $MCD_{95\%}$, estimado a través del error estándar de medición, de 5,5, fue similar al obtenido con el cuestionario en su versión original ($MCD_{90\%}$ 4,85 puntos) (380), y es inferior a la MDCl estimada a través de las preguntas transicionales, lo que da congruencia a los resultados.

Se han hecho algunas comparaciones, a nivel internacional entre el OHS y otras herramientas de evaluación de la CVRS que son ampliamente utilizadas en los pacientes con artrosis de cadera.

El OHS se comparó con la escala de Harris (Harris Hip Score, HHS) para valorar si este instrumento puede reemplazar al HHS para la evaluación de los pacientes sometidos a cirugía de reemplazo articular (383). El HHS es uno de los instrumentos más utilizados para evaluar los resultados obtenidos tras la artroplastia de cadera. En el análisis de la fiabilidad, la consistencia interna del OHS fue notablemente superior a la HHS en todas las ocasiones. La capacidad discriminativa del cuestionario fue superior para el HHS, comparado con el OHS, con un tamaño del efecto mayor de 4,1 frente al 2,1. Ambos cuestionarios mostraron efecto techo después de la cirugía, pero la OHS en menor medida (36,4%) que el HHS (55,6%). En general, los resultados tras la cirugía de reemplazo articular son satisfactorios, por lo que los cuestionarios con efectos de techo bajo son favorables para comparar los resultados entre los pacientes. Se analizó la validez convergente mediante la correlación de las escala del HHS y OHS al inicio del estudio y un año después de la operación. Mostrando una correlación moderadamente alta entre las puntuaciones resumen al inicio y un año después de la operación, coeficiente de correlación de Spearman, 0,57 basal y 0,65 al año. Comparando los resultados se puede considerar el OHS como un instrumento válido para la evaluación de la CVRS después de la cirugía de reemplazo total de cadera (383).

La correlación de las escalas del HHS y OHS es moderadamente alta al año de la cirugía de cadera, pero existe la preocupación de que la correlación no persista a corto o medio plazo, fundamentalmente teniendo en cuenta que los problemas de cadera son generalmente poco frecuentes en este periodo. En este sentido, encontramos estudios que comparan ambos índices, 5 años después de la cirugía de artroplastia de cadera (252), donde se muestra al OHS como una herramienta comparable al complejo sistema de puntuación del HHS, encontrándose correlaciones altas entre ambos (coeficiente de correlación de Spearman 0,712).

Otro estudio comparativo reciente, pretendió determinar si tanto el OHS como WOMAC, son instrumentos válidos para predecir una mejoría clínica significativa, después de una cirugía de

artroplastia total de cadera (384). Los resultados concluyen que tanto la puntuación preoperatoria del OHS, como del cuestionario WOMAC, pueden ser útiles para predecir una mejoría clínicamente significativa, pero sin embargo, su capacidad para predecir la satisfacción del paciente es limitada. No obstante, está bien establecido que la satisfacción del paciente está influenciada por factores no directamente relacionados con la intervención quirúrgica, como la experiencia de la estancia hospitalaria o el nivel socioeconómico (385,386), por lo que un resultado clínico aparentemente bueno no predice necesariamente un paciente satisfecho.

Otros estudios han determinado que la puntuación prequirúrgica del OHS puede ser útil para “predecir” el éxito quirúrgico de la artroplastia, desde la perspectiva del paciente (387). Tanto las puntuaciones de OHS como de WOMAC se pueden utilizar como predictores de mejoría clínica significativa, proporcionando información útil sobre las mejoras esperadas después de la cirugía de cadera para pacientes y cirujanos. Sin embargo, las puntuaciones preoperatorias no se pueden utilizar para predecir directamente la satisfacción.

Queda pues establecido que el OHS es una herramienta con capacidades discriminativas y evaluativas adecuadas.

7.2.1. Limitaciones

Este trabajo tiene algunas limitaciones. La población no tiene por qué ser representativa de la población nacional, aunque se ha incluido a un grupo de pacientes de diversas zonas y con diferentes estados evolutivos de su enfermedad. Por otra parte, tiene algunas de las limitaciones propias de la metodología empleada para medir las propiedades psicométricas, la teoría clásica de los test (TCT), con su carga de presunciones y matices, pero se ha complementado el proceso de validación con el AFC específico para datos categóricos, que somete sus asunciones apriorísticas a contrastes estadísticos (351).

7.2.2. Implicaciones

Tradicionalmente, el OHS se ha utilizado para evaluar el impacto de la cirugía de prótesis de cadera (388) en la CVRS, así como de otros procedimientos quirúrgicos (389) o no quirúrgicos (256). Pocos estudios se han centrado en su capacidad discriminatoria (258,263). En este trabajo

DISCUSIÓN

se destaca la capacidad discriminatoria de la herramienta y se valora su sensibilidad a los cambios en la evolución general de la enfermedad, incluyendo tanto a los pacientes que se someten a una prótesis articular como a los que no. Su utilidad es similar a la de otros instrumentos con amplia experiencia de uso tras ser adaptados a otros idiomas, como el WOMAC, y muestra mayor capacidad para detectar cambios que herramientas genéricas como el EQ-5D-5L.

7.3. Validación del Oxford Knee Score (OKS)

El cuestionario OKS en su versión en castellano es una herramienta válida para medir la CVRS en pacientes con OA de rodilla, es fiable y sensible al cambio. Además es una herramienta aceptable por la población a la que se dirige, pues la tasa de respuestas es extraordinariamente alta.

La validez del OKS se ha valorado desde varias perspectivas. No se estudió su validez “aparente” al tratarse de una adaptación. La validez convergente o de grupos conocidos parece adecuada, pues la puntuación obtenida es muy diferente en sujetos con puntuaciones diferentes en las escalas del WOMAC o en el EQ-5D-5L. Además no parece tener efectos techo y suelo relevantes que comprometan dicha capacidad discriminativa, como ya se había señalado en otras adaptaciones (303,304).

La validez convergente de la herramienta pareció correcta. Las correlaciones con las escalas específicas del WOMAC y genéricas del EQ-5D-5L fueron más fuertes que las encontradas en la versión original con herramientas genéricas de CVRS, aunque en este caso se usó el SF-36 (296), y similares o ligeramente superiores a las encontrada para adaptaciones del OKS a otros idiomas como el portugués (306), el alemán (303), italiano (300), holandés (301), persa (308), coreano (307), francés (304), finlandés (311), griego (309), japonés (305) y árabe (310). La versión al español de Colombia (263) no presenta datos relativos a la validez convergente de la herramienta.

Las correlaciones de las puntuaciones del OKS y el EQ-5D-5L fueron también más fuertes en nuestro caso que los obtenidos en estudios realizados en población inglesa (298). Esta forma de medir la validez convergente ofrece mejores resultados para el OKS que las reportadas para otros cuestionarios como el KSS, en su versión adaptada al castellano (390).

Respecto a la consistencia interna, el alfa de Cronbach fue mejor que el encontrado para la escala original en el periodo de inclusión (0,92 vs 0,87) (296). Como se comentó para el caso del OHS, un valor muy alto del el alfa de Cronbach podría indicar redundancia de los ítems, situación descartable tras la revisión del análisis factorial.

La fiabilidad test-retest, medida mediante el CCI, fue elevada, y, aunque esta forma de medir la fiabilidad suele sobrestimarla (391), los valores obtenidos superiores a 0,90, en una muestra de más de 150 pacientes permiten calificar a la herramienta como fiable (323).

DISCUSIÓN

El estudio de la fiabilidad de esta versión adaptada del OKS arrojó valores para el alfa de Cronbach y para la fiabilidad test-retest superiores a las obtenidas para otras herramientas construidas con el mismo fin, y que también son ampliamente utilizadas en lengua inglesa como el Knee Injury and Osteoarthritis Outcome Score (KOOS) (392).

El otro elemento estudiado respecto a su fiabilidad fue la dimensionalidad. Su estructura factorial, confirma la unidimensionalidad de la versión evaluada del cuestionario. En el AFE se encontró que todos los ítems saturaban de manera consistente en el mismo factor con valores superiores a los de la versión inglesa (J Dawson et al., 1998). En el AFC se obtuvieron valores del RMSEA aceptables, con índices TLI y CFI óptimos (Schreiber et al., 2006), que invitan a pensar en que la estructura factorial propuesta era la correcta.

Otros autores han estudiado la posibilidad de desagregar los componentes de dolor y discapacidad del OKS (393). Los hallazgos de otras investigaciones realizadas sobre este tema sugieren que el dolor y la función autoinformados son constructos que pueden tener una superposición considerable, particularmente en la forma en que los pacientes los perciben, esto se debe a que el dolor influye en la capacidad funcional (394–396).

La existencia de una estructura bidimensional, al igual que comentamos para el OHS permitiría extraer información de la dimensión dolor y función en forma de subescalas por separado en lugar de un único valor resumen de CVRS. En este sentido se postula que algunos ítems saturan de manera consistente en los factores denominados función (ítems 2, 3, 7, 11 y 12) y dolor (ítems 1, 6 y 8), mientras que los ítems de carga cruzada (ítems 4, 5, 9 y 10) que preguntan sobre el dolor en relación con la función, se asignarían al dominio dolor del OKS. No obstante, el estudio de valoración de la dimensionalidad de las escalas no demuestra la superioridad de este modelo y determina que ambos modelos son plausibles (393). Nuestros resultados hacen elegible, en esta adaptación la versión unidimensional.

Se debe reseñar la aplicabilidad descrita de la versión de dos subescalas. La subescala de dolor del Oxford Knee Score (OKS-PS) se ha utilizado para identificar a los pacientes con dolor crónico después de una artroplastia de rodilla (397). Emplea una puntuación resumida de la subescala OKS-PS sumando respuestas de sus siete ítems, por lo que van desde 0 (más dolor) a 28 (menos dolor). Se determinó que un valor de corte con una puntuación del OKS-PS ≤ 14 a los seis meses de la cirugía que podría considerarse dolor crónico, y que probablemente afectara negativamente a la calidad de vida relacionada con la salud.

El OKS se comparó con el cuestionario KOOS para evaluar la respuesta al cambio de pacientes intervenidos de rodilla (398). Todos los instrumentos demostraron cambios estadísticamente significativos entre las evaluaciones preoperatorias y al año de la cirugía. Todos los cambios mostraron una mejora en la puntuación, pero los instrumentos específicos revelaron una mayor capacidad de respuesta que los cuestionarios genéricos de CVRS. Tanto el OKS, como todas las subescalas del KOOS mostraron puntuaciones de capacidad de respuesta al cambio medidas mediante el tamaño del efecto mayores a 1,24, muy superiores al valor de 0,8 que determina el umbral para determinar cambios grandes en la CVRS (357).

La capacidad discriminativa de la adaptación del cuestionario, la que hace referencia a su potencial para distinguir entre sujetos con situaciones diferentes, es adecuada, pero, como ocurría con el OHS, la herramienta también puede ser utilizada para estudiar el cambio de percepción respecto a su situación de un mismo sujeto, esto es, tiene una adecuada capacidad evaluativa (147). De hecho esta herramienta se diseñó para tal fin y los resultados obtenidos avalan este uso.

El tamaño del efecto para cambios positivos “moderados” estuvo un poco por debajo del umbral de 0,8, en valores similares al WOMAC. El tamaño del efecto detectado para los cambios en el trabajo de validación original tras cirugía de 2,1 (296) debe compararse con las mejoras muy importantes (tamaño del efecto 1,6) pues se testó en pacientes que habían sido intervenidos de la rodilla.

La capacidad evaluativa, como en otros cuestionarios, es superior para detectar cambios positivos que negativos (382), aunque incluso en los cambios negativos tiene una capacidad similar o superior al WOMAC y, por supuesto, al cuestionario genérico EQ-5D-5L.

La MDCl para los sujetos que notan mejoría moderada fue de 6,1 puntos. Atendiendo al criterio de estimar la MDCl a través de valorar la mitad de una desviación estándar de la distribución de las puntuaciones en los sujetos que cambian (381), podríamos situar la MDCl alrededor de los 4,5 puntos. Valores similares para la MDCl (entre 3 y 5 puntos) se habían propuesto en la validación de la escala original (250) y se han confirmado en trabajos posteriores (399).

El $MCD_{95\%}$, estimado a través del error estándar de medición, de 4,38, es inferior a la MDCl estimada a través de las preguntas transicionales, lo que da congruencia a los resultados.

7.3.1. Limitaciones

Este trabajo tiene algunas limitaciones. La población no se eligió con criterios de representatividad de la población nacional, aunque se ha incluido a un grupo de pacientes de diversas zonas y con diferentes estados evolutivos de su enfermedad. Por otra parte, tiene algunas de las limitaciones propias de la metodología empleada para medir las propiedades psicométricas, la teoría clásica de los test (TCT), por lo que el proceso de validación también se ha complementado con el AFC específico para datos categóricos, que somete sus asunciones a contrastes estadísticos (351).

7.3.2 Implicaciones

Los resultados aportados permiten proponer el uso del OKS para aquellas situaciones en las que se ha utilizado su versión original, como la mejora de la CVRS tras la cirugía de remplazo de rodilla (298), o el estudio de los factores asociados a los resultados de la cirugía (337,400,401), pero también para discriminar entre pacientes con diferentes situaciones clínica y para valorar la evolución de esta situación en el tiempo, debido a su capacidad para detectar mejoras “moderadas” en los pacientes.

Una vez discutido el impacto de la artrosis de miembro inferior sobre la CVRS y la utilidad de las herramientas adaptadas para su valoración, se pasa a concluir sobre los objetivos de este trabajo.

Caracterización de la Calidad de Vida Relacionada con la Salud en pacientes con artrosis de miembro inferior y validación de nuevas herramientas para su evaluación.

DISCUSIÓN

8. Conclusiones

CONCLUSIONES

1. La artrosis de miembros inferiores (OAMI) se asocia a una enorme disminución en la utilidad atribuida al estado de salud, una vez excluida la influencia de otras variables que se relacionan con la Calidad de Vida Relacionada con la Salud (CVRS). El impacto producido por la OAMI sobre la CVRS medida en utilidades supera con mucho los umbrales establecidos para la relevancia clínica.
2. La artrosis de rodilla empeora la percepción de CVRS en menor medida que la artrosis de cadera, una vez ajustada esta asociación por otras potenciales características que actúan como variables de confusión. Estas diferencias son cercanas a los umbrales a los que se atribuye relevancia clínica en la literatura.
3. Cuando se comparan con la población general, las personas con artrosis de miembro inferior de mayor edad, las de grupos socioeconómicos más desfavorecidos, las que viven solas, las mujeres, las personas obesas y las que presentan más condiciones crónicas, suelen tener peor percepción del estado de salud. Las diferencias en percepción de CVRS entre personas ya diagnosticadas de artrosis de miembro inferior solo vienen explicadas por el tipo de artrosis, el sexo (peor percepción las mujeres), la presencia de otras enfermedades
4. El cuestionario Oxford Hip Score (OHS), en su versión en castellano, es una herramienta aceptable válida y fiable para medir la CVRS en pacientes con artrosis de cadera.
5. La capacidad de la versión en castellano del OHS para evaluar cambios con relevancia clínica a nivel individual, en sujetos con artrosis de cadera, es adecuada, distinguiendo con claridad los cambios potenciales del “ruido aleatorio” propio de cualquier instrumento de medida.

CONCLUSIONES

6. El cuestionario Oxford Knee Score (OKS), en su versión en castellano, es una herramienta aceptable, válida y fiable para medir la CVRS en pacientes con artrosis de rodilla.

7. La capacidad de la versión en castellano del OKS para evaluar cambios con relevancia clínica a nivel individual, en sujetos con artrosis de rodilla, es adecuada, permitiendo distinguir con claridad los cambios potenciales del “ruido aleatorio” propio de cualquier instrumento de medida.

9. Bibliografía

BIBLIOGRAFÍA

1. Garriga XM. Definición, etiopatogenia, clasificación y formas de presentación. *Atención Primaria*. 2014 Jan;46:3–10.
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. *Osteoarthr Cartil*. 2015;23(8):1233–41.
3. Blackburn S, Research User Group, Rhodes C, Higginbottom A, Dziedzic K. The OARSI standardised definition of osteoarthritis: A lay version. *Osteoarthr Cartil*. 2016 Apr;24:S192.
4. van den Bosch MH, Blom AB, Kram V, Maeda A, Sikka S, Gabet Y, et al. WISP1/CCN4 aggravates cartilage degeneration in experimental osteoarthritis. *Osteoarthr Cartil*. 2017 Nov;25(11):1900–11.
5. Martijn Hvan den Bosch., Blom AB, Sloetjes AW, Koenders MI, van de Loo FA, van den Berg WB, et al. Induction of Canonical Wnt Signaling by Synovial Overexpression of Selected Wnts Leads to Protease Activity and Early Osteoarthritis-Like Cartilage Damage. *Am J Pathol*. 2015 Jul;185(7):1970–80.
6. Ghoochani N, Karandish M, Mowla K, Haghhighizadeh MH, Jalali MT. The effect of pomegranate juice on clinical signs, matrix metalloproteinases and antioxidant status in patients with knee osteoarthritis. *J Sci Food Agric*. 2016 Oct;96(13):4377–81.
7. Lin PM, Chen C-TC, Torzilli PA. Increased stromelysin-1 (MMP-3), proteoglycan degradation (3B3- and 7D4) and collagen damage in cyclically load-injured articular cartilage. *Osteoarthr Cartil*. 2004 Jun;12(6):485–96.
8. Wang CC, Guo L, Tian FD, An N, Luo L, Hao RH, et al. Naringenin regulates production of matrix metalloproteinases in the knee-joint and primary cultured articular chondrocytes and alleviates pain in rat osteoarthritis model. *Brazilian J Med Biol Res*. 2017;50(4).
9. Bai Y, Gao S, Liu Y, Jin S, Zhang H, Su K. Correlation between Interleukin-17 gene polymorphism and osteoarthritis susceptibility in Han Chinese population. *BMC Med Genet*. 2019 Dec 18;20(1):20.
10. Özler K, Aktaş E, Atay Ç, Yılmaz B, Arıkan M, Güngör Ş. Serum and knee synovial fluid matrixmetalloproteinase-13 and tumor necrosis factor-alpha levels in patients with late stage osteoarthritis. *Acta Orthop Traumatol Turc*. 2016 Dec;50(6):670–3.
11. Brody LT. Knee osteoarthritis: Clinical connections to articular cartilage structure and function. *Phys Ther Sport*. 2015 Nov;16(4):301–16.
12. Litwic A, Edwards MH, Dennison EM, Cooper C. Epidemiology and burden of osteoarthritis. *Br Med Bull*. 2013 Mar 1;105(1):185–99.
13. Deveza LA, Nelson AE, Loeser RF. Phenotypes of osteoarthritis: current state and future implications. *Clin Exp Rheumatol*. 2019;37(5):64–72.
14. Anderson GP. Endotyping asthma: new insights into key pathogenic mechanisms in a complex, heterogeneous disease. *Lancet*. 2008;372(9643):1107–19.

BIBLIOGRAFÍA

15. Arden N, Nevitt M. Osteoarthritis: Epidemiology. *Best Pract Res Clin Rheumatol*. 2006 Feb;20(1):3–25.
16. Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan JL, Protheroe J, Jordan KP. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. *Osteoarthr Cartil*. 2015 Apr;23(4):507–15.
17. Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. *Osteoarthr Cartil*. 2005;13(9):769–81.
18. Zhang Y, Hunter DJ, Nevitt MC, Xu L, Niu J, Lui L-Y, et al. Association of squatting with increased prevalence of radiographic tibiofemoral knee osteoarthritis: The Beijing Osteoarthritis Study. *Arthritis Rheum*. 2004 Apr;50(4):1187–92.
19. Nelson AE, Braga L, Renner JB, Atashili J, Woodard J, Hochberg MC, et al. Characterization of individual radiographic features of hip osteoarthritis in African American and white women and men: The Johnston County Osteoarthritis Project. *Arthritis Care Res*. 2010;62(2):190–7.
20. Reyes C, Leyland KM, Peat G, Cooper C, Arden NK, Prieto-Alhambra D. Association Between Overweight and Obesity and Risk of Clinically Diagnosed Knee, Hip, and Hand Osteoarthritis: A Population-Based Cohort Study. *Arthritis Rheumatol*. 2016;68(8):1869–75.
21. Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Obesity and osteoarthritis in knee, hip and/or hand: An epidemiological study in the general population with 10 years follow-up. *BMC Musculoskelet Disord*. 2008;9:1–5.
22. Felson DT, Zhang Y. Smoking and osteoarthritis: A review of the evidence and its implications. *Osteoarthr Cartil*. 2015;23(3):331–3.
23. Hui M, Doherty M, Zhang W. Does smoking protect against osteoarthritis? Meta-analysis of observational studies. *Ann Rheum Dis*. 2011;70(7):1231–7.
24. Zhang FF, Driban JB, Lo GH, Price LL, Booth S, Eaton CB, et al. 2014-F.Zhang-Vitamin D Deficiency is Associated with progression of knee OA. 2014;
25. Laslett LL, Quinn S, Burgess JR, Parameswaran V, Winzenberg TM, Jones G, et al. Moderate Vitamin D deficiency is associated with changes in knee and hip pain in older adults: A 5-year Longitudinal study. *Ann Rheum Dis*. 2014;73(4):697–703.
26. Joseph GB, Jan, Lynch JA, Lane NE, Link TM. Associations Between Vitamins C and D Intake and Cartilage Composition and Knee Joint Morphology Over 4 Years: Data From the Osteoarthritis Initiative McCulloch, Charles E. Nevitt, Michael C. Neumann. *Arthritis Care Res (Hoboken)*. 2020 Sep;72(9):1239–47.
27. Chin K-Y. The Relationship between Vitamin K and Osteoarthritis: A Review of Current Evidence. *Nutrients*. 2020 Apr 25;12(5):1208.
28. Chin KY, Ima-Nirwana S. The role of vitamin E in preventing and treating osteoarthritis - A review of the current evidence. *Front Pharmacol*. 2018;9(AUG):1–14.
29. Sharma G, Saxena RK, Mishra P. Regeneration of static-load-degenerated articular cartilage extracellular matrix by vitamin C supplementation. *Cell Tissue Res*. 2008 Oct 5;334(1):111–20.

30. Chaganti RK, Tolstykh I, Javaid MK, Neogi T, Torner J, Curtis J, et al. High plasma levels of vitamin C and E are associated with incident radiographic knee osteoarthritis. *Osteoarthritis Cartil.* 2014;22(2):190–6.
31. Calder PC. n-3 Polyunsaturated fatty acids, inflammation, and inflammatory diseases. *Am J Clin Nutr.* 2006 Jun 1;83(6):1505S-1519S.
32. Baker KR, Et A. Ass'n of Plasma n-6 & n-3 PUFAs w synovitis in the knee: the MOST study. 2013;20(5):382–7.
33. Loef M, Schoones JW, Kloppenburg M, Ioan-Facsinay A. Fatty acids and osteoarthritis: different types, different effects. *Jt Bone Spine.* 2019 Jul;86(4):451–8.
34. Koeberle A, Werz O. Multi-target approach for natural products in inflammation. *Drug Discov Today.* 2014 Dec;19(12):1871–82.
35. He Y, Yue Y, Zheng X, Zhang K, Chen S, Du Z. Curcumin, Inflammation, and Chronic Diseases: How Are They Linked? *Molecules.* 2015 May 20;20(5):9183–213.
36. Daily JW, Yang M, Park S. Efficacy of Turmeric Extracts and Curcumin for Alleviating the Symptoms of Joint Arthritis: A Systematic Review and Meta-Analysis of Randomized Clinical Trials. *J Med Food.* 2016;19(8):717–29.
37. Rodriguez-Fontenla C, Gonzalez A. Genética de la artrosis. *Reumatol Clínica.* 2015 Jan;11(1):33–40.
38. Yucesoy B, Charles LE, Baker B, Burchfiel CM. Occupational and genetic risk factors for osteoarthritis: A review. *Work.* 2015;50(2):261–73.
39. Zhang M, Wang J. Epigenetic regulation of gene expression in osteoarthritis. *Genes Dis.* 2015;2(1):69–75.
40. Miyamoto Y, Mabuchi A, Shi D, Kubo T, Takatori Y, Saito S, et al. A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis. *Nat Genet.* 2007 Apr 25;39(4):529–33.
41. Kerkhof HJM, Lories RJ, Meulenbelt I, Jonsdottir I, Valdes AM, Arp P, et al. A genome-wide association study identifies a locus on chromosome 7q22 to influence susceptibility for osteoarthritis. *Arthritis Rheum.* 2010;NA-NA.
42. Yau MS, Yerges-Armstrong LM, Liu Y, Lewis CE, Duggan DJ, Renner JB, et al. Genome-Wide Association Study of Radiographic Knee Osteoarthritis in North American Caucasians. *Arthritis Rheumatol.* 2017 Feb;69(2):343–51.
43. Warner SC, Valdes AM. Genetic association studies in osteoarthritis. *Curr Opin Rheumatol.* 2017 Jan;29(1):103–9.
44. Lotz M, Martel-Pelletier J, Christiansen C, Brandi ML, Bruyère O, Chapurlat R, et al. Value of biomarkers in osteoarthritis: Current status and perspectives. *Postgrad Med J.* 2014;90(1061):171–8.
45. Valdes AM, Meulenbelt I, Chassaing E, Arden NK, Bierma-Zeinstra S, Hart D, et al. Large scale meta-analysis of urinary C-terminal telopeptide, serum cartilage oligomeric protein and matrix metalloprotease degraded type II collagen and their role in prevalence, incidence and progression of osteoarthritis. *Osteoarthritis Cartil.* 2014;22(5):683–9.

BIBLIOGRAFÍA

46. Saberi Hosnijeh F, Runhaar J, Van Meurs JBJ, Bierma-Zeinstra SM. Biomarkers for osteoarthritis: Can they be used for risk assessment? A systematic review. *Maturitas*. 2015;82(1):36–49.
47. Mazières B, Garnero P, Guéguen A, Abbal M, Berdah L, Lequesne M, et al. Molecular markers of cartilage breakdown and synovitis at baseline as predictors of structural progression of hip osteoarthritis. The ECHODIAH* cohort. *Ann Rheum Dis*. 2006;65(3):354–9.
48. Huebner JL, Bay-Jensen AC, Huffman KM, He Y, Leeming DJ, McDaniel GE, et al. Alpha C-telopeptide of type I collagen is associated with subchondral bone turnover and predicts progression of joint space narrowing and osteophytes in osteoarthritis. *Arthritis Rheumatol*. 2014;66(9):2440–9.
49. Blumenfeld O, Williams FMK, Hart DJ, Spector TD, Arden N, Livshits G. Association between cartilage and bone biomarkers and incidence of radiographic knee osteoarthritis (RKO) in UK females: a prospective study. *Osteoarthr Cartil*. 2013 Jul;21(7):923–9.
50. Golightly YM, Marshall SW, Kraus VB, Renner JB, Villaveces A, Casteel C, et al. Biomarkers of incident radiographic knee osteoarthritis: Do they vary by chronic knee symptoms? *Arthritis Rheum*. 2011 Aug;63(8):2276–83.
51. Kelman A, Lui L, Yao W, Krumme A, Nevitt M, Lane NE. Association of higher levels of serum cartilage oligomeric matrix protein and N-telopeptide crosslinks with the development of radiographic hip osteoarthritis in elderly women. *Arthritis Rheum*. 2006 Jan;54(1):236–43.
52. Sharif M. Increased serum C reactive protein may reflect events that precede radiographic progression in osteoarthritis of the knee. *Ann Rheum Dis*. 2000 Jan 1;59(1):71–4.
53. Sharif M, George E, Shepstone L, Knudson W, Thonar EJ-. A, Cushnaghan J, et al. Serum hyaluronic acid level as a predictor of disease progression in osteoarthritis of the knee. *Arthritis Rheum*. 1995;38(6):760–7.
54. Kraus VB, Hargrove DE, Hunter DJ, Renner JB, Jordan JM. Establishment of reference intervals for osteoarthritis-related soluble biomarkers: the FNIH/OARSI OA Biomarkers Consortium. *Ann Rheum Dis*. 2017 Jan;76(1):179–85.
55. van Spil WE, DeGroot J, Lems WF, Oostveen JCM, Lafeber FPJG. Serum and urinary biochemical markers for knee and hip-osteoarthritis: A systematic review applying the consensus BIPED criteria. *Osteoarthr Cartil*. 2010;18(5):605–12.
56. Lafeber FPJG, van Spil WE. Osteoarthritis year 2013 in review: Biomarkers; reflecting before moving forward, one step at a time. *Osteoarthr Cartil*. 2013;21(10):1452–64.
57. Kraus VB. Osteoarthritis year 2010 in review: Biochemical markers. *Osteoarthr Cartil*. 2011;19(4):346–53.
58. Ren G, Krawetz RJ. Biochemical Markers for the Early Identification of Osteoarthritis: Systematic Review and Meta-Analysis. *Mol Diagn Ther*. 2018 Dec 30;22(6):671–82.
59. Bauer DC, Hunter DJ, Abramson SB, Attur M, Corr M, Felson D, et al. Classification of osteoarthritis biomarkers: a proposed approach. *Osteoarthr Cartil*. 2006 Aug;14(8):723–7.
60. Hardcastle SA, Dieppe P, Gregson CL, Arden NK, Spector TD, Hart DJ, et al. Individuals with high bone mass have an increased prevalence of radiographic knee osteoarthritis. *Bone*. 2015;71:171–9.

61. Hardcastle SA, Dieppe P, Gregson CL, Hunter D, Thomas GER, Arden NK, et al. Prevalence of radiographic hip osteoarthritis is increased in high bone mass. *Osteoarthr Cartil.* 2014;22(8):1120–8.
62. Puenpatom RA, Victor TW. Increased Prevalence of Metabolic Syndrome in Individuals with Osteoarthritis: An Analysis of NHANES III Data. *Postgrad Med.* 2009 Nov 1;121(6):9–20.
63. Courties A, Sellam J, Berenbaum F. Metabolic syndrome-associated osteoarthritis. *Curr Opin Rheumatol.* 2017 Mar;29(2):214–22.
64. Courties A, Gualillo O, Berenbaum F, Sellam J. Metabolic stress-induced joint inflammation and osteoarthritis. *Osteoarthr Cartil.* 2015;23(11):1955–65.
65. Vega-Franco L, Zhuo Q, Yang W, Chen J, Wang Y. Metabolic syndrome meets osteoarthritis. *Nat Rev Rheumatol.* 2012 Dec 21;8(12):729–37.
66. de Munter W, van der Kraan PM, van den Berg WB, van Lent PLEM. High systemic levels of low-density lipoprotein cholesterol: fuel to the flames in inflammatory osteoarthritis? *Rheumatology (Oxford).* 2016 Jan;55(1):16–24.
67. Antony B, Venn A, Cicuttini F, March L, Blizzard L, Dwyer T, et al. Correlates of knee bone marrow lesions in younger adults. *Arthritis Res Ther.* 2016;18(1):1–9.
68. Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, et al. Mutual associations among musculoskeletal diseases and metabolic syndrome components: A 3-year follow-up of the ROAD study. *Mod Rheumatol.* 2015 May 4;25(3):438–48.
69. Juhakoski R, Heliövaara M, Impivaara O, Kröger H, Knekt P, Lauren H, et al. Risk factors for the development of hip osteoarthritis: a population-based prospective study. *Rheumatology (Oxford).* 2009 Jan;48(1):83–7.
70. Ajuied A, Wong F, Smith C, Norris M, Earnshaw P, Back D, et al. Anterior Cruciate Ligament Injury and Radiologic Progression of Knee Osteoarthritis. *Am J Sports Med.* 2014 Sep 8;42(9):2242–52.
71. Kopf S, Forsythe B, Wong AK, Tashman S, Anderst W, Irrgang JJ, et al. Nonanatomic Tunnel Position in Traditional Transtibial Single-Bundle Anterior Cruciate Ligament Reconstruction Evaluated by Three-Dimensional Computed Tomography. *J Bone Jt Surgery-American Vol.* 2010 Jun;92(6):1427–31.
72. Kopf S, Forsythe B, Wong AK, Tashman S, Irrgang JJ, Fu FH. Transtibial ACL reconstruction technique fails to position drill tunnels anatomically in vivo 3D CT study. *Knee Surgery, Sport Traumatol Arthrosc.* 2012 Nov;20(11):2200–7.
73. Liu C, Wang Y, Li Z, Li J, Zhang H, Fu Y, et al. Tibiofemoral joint contact area and stress after single-bundle anterior cruciate ligament reconstruction with transtibial versus anteromedial portal drilling techniques. *J Orthop Surg Res.* 2018 Dec 4;13(1):247.
74. Abebe ES, Utturkar GM, Taylor DC, Spritzer CE, Kim JP, Moorman CT, et al. The effects of femoral graft placement on in vivo knee kinematics after anterior cruciate ligament reconstruction. *J Biomech.* 2011 Mar;44(5):924–9.
75. Rothrauff BB, Jorge A, de Sa D, Kay J, Fu FH, Musahl V. Anatomic ACL reconstruction reduces risk of post-traumatic osteoarthritis: a systematic review with minimum 10-year

- follow-up. *Knee Surgery, Sport Traumatol Arthrosc.* 2020;28(4):1072–84.
76. Poulsen E, Goncalves GH, Bricca A, Roos EM, Thorlund JB, Juhl CB. Knee osteoarthritis risk is increased 4-6 fold after knee injury-a systematic review and meta-analysis. *Br J Sports Med.* 2019;53(23):1454–63.
77. Papalia R, Del Buono A, Osti L, Denaro V, Maffulli N. Meniscectomy as a risk factor for knee osteoarthritis: a systematic review. *Br Med Bull.* 2011 Sep 1;99(1):89–106.
78. Roemer FW, Kwok CK, Hannon MJ, Hunter DJ, Eckstein F, Grago J, et al. Partial meniscectomy is associated with increased risk of incident radiographic osteoarthritis and worsening cartilage damage in the following year. *Eur Radiol.* 2017 Jan 27;27(1):404–13.
79. Schram B, Orr R, Pope R, Canetti E, Knapik J. Risk factors for development of lower limb osteoarthritis in physically demanding occupations: A narrative umbrella review. *J Occup Health.* 2020;62(1):1–13.
80. Sandmark H, Hogstedt C, Vingård E. Primary osteoarthrosis of the knee in men and women as a result of lifelong physical load from work. *Scand J Work Environ Health.* 2000 Feb;26(1):20–5.
81. Kaila-Kangas L, Arokoski J, Impivaara O, Viikari-Juntura E, Leino-Arjas P, Luukkonen R, et al. Associations of hip osteoarthritis with history of recurrent exposure to manual handling of loads over 20 kg and work participation: A population-based study of men and women. *Occup Environ Med.* 2011;68(10):734–8.
82. Allen KD, Chen J-C, Callahan LF, Golightly YM, Helmick CG, Renner JB, et al. Associations of Occupational Tasks with Knee and Hip Osteoarthritis: The Johnston County Osteoarthritis Project. *J Rheumatol.* 2010 Apr;37(4):842–50.
83. Yoshimura N, Sasaki S, Iwasaki K, Danjoh S, Kinoshita H, Yasuda T, et al. Occupational lifting is associated with hip osteoarthritis: a Japanese case-control study. *J Rheumatol.* 2000 Feb;27(2):434–40.
84. Neogi T, Zhang Y. Epidemiology of Osteoarthritis. *Rheum Dis Clin North Am.* 2013 Feb;39(1):1–19.
85. Alentorn-Geli E, Samuelsson K, Musahl V, Green CL, Bhandari M, Karlsson J. The Association of Recreational and Competitive Running With Hip and Knee Osteoarthritis: A Systematic Review and Meta-analysis. *J Orthop Sport Phys Ther.* 2017 Jun;47(6):373–90.
86. Vannini F, Spalding T, Andriolo L, Berruto M, Denti M, Espregueira-Mendes J, et al. Sport and early osteoarthritis: the role of sport in aetiology, progression and treatment of knee osteoarthritis. *Knee Surgery, Sport Traumatol Arthrosc.* 2016 Jun 4;24(6):1786–96.
87. Sharma L, Song J, Dunlop D, Felson D, Lewis CE, Segal N, et al. Varus and valgus alignment and incident and progressive knee osteoarthritis. *Ann Rheum Dis.* 2010 Nov;69(11):1940–5.
88. Tanamas S, Hanna FS, Cicuttini FM, Wluka AE, Berry P, Urquhart DM. Does knee malalignment increase the risk of development and progression of knee osteoarthritis? A systematic review. *Arthritis Rheum.* 2009 Apr 15;61(4):459–67.
89. Felson DT, Niu J, Gross KD, Englund M, Sharma L, Cooke TD V., et al. Valgus malalignment is a risk factor for lateral knee osteoarthritis incidence and progression: Findings from the multicenter osteoarthritis study and the osteoarthritis initiative. *Arthritis Rheum.*

2013;65(2):355–62.

90. Wei J, Gross D, Lane NE, Lu N, Wang M, Zeng C, et al. Risk factor heterogeneity for medial and lateral compartment knee osteoarthritis: analysis of two prospective cohorts. *Osteoarthr Cartil.* 2019 Apr;27(4):603–10.
91. Zeng WN, Wang FY, Chen C, Zhang Y, Gong XY, Zhou K, et al. Investigation of association between hip morphology and prevalence of osteoarthritis. *Sci Rep.* 2016;6(March):1–8.
92. Saberi Hosnijeh F, Zuiderwijk ME, Versteeg M, Smeele HTW, Hofman A, Uitterlinden AG, et al. Cam Deformity and Acetabular Dysplasia as Risk Factors for Hip Osteoarthritis. *Arthritis Rheumatol.* 2017 Jan 2;69(1):86–93.
93. Kellgren JH, Lawrence JS. Radiological Assessment of Osteo-Arthrosis. *Ann Rheum Dis.* 1957 Dec 1;16(4):494–502.
94. 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. *Arthritis Rheum.* 1986;29(8):1039–49.
95. Altman R, Alarcón G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum.* 1991 May;34(5):505–14.
96. Altman R, Alarcón G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. *Arthritis Rheum.* 1990 Nov;33(11):1601–10.
97. Murray CJ., Lopez AD. The Global burden of disease : a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020 : summary. 1996.
98. Comas M, Sala M, Román R, Hoffmeister L, Castells X. Variaciones en la estimación de la prevalencia de artrosis de rodilla según los criterios diagnósticos utilizados en los estudios poblacionales. *Gac Sanit.* 2010;24(1):28–32.
99. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet.* 2018;392(10159):1789–858.
100. Zhao X, Shah D, Gandhi K, Wei W, Dwibedi N, Webster L, et al. Clinical, humanistic, and economic burden of osteoarthritis among noninstitutionalized adults in the United States. *Osteoarthr Cartil.* 2019 Nov;27(11):1618–26.
101. Blanco FJ, Silva-Díaz M, Quevedo Vila V, Seoane-Mato D, Pérez Ruiz F, Juan-Mas A, et al. Prevalencia de artrosis sintomática en España: Estudio EPISER2016. *Reumatol Clínica.* 2020 Apr;(xx):1–10.
102. Puig-Junoy J, Ortún-Rubio V, Pinto-Prades JL. Los costes en la evaluación económica de tecnologías sanitarias. *Atención Primaria.* 2001;27(3):186–9.
103. Vos T, Allen C, Arora M, Barber RM, Brown A, Carter A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet.* 2016

BIBLIOGRAFÍA

- Oct;388(10053):1545–602.
104. Vos T, Allen C, Arora M, Barber RM, Brown A, Carter A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*. 2016;388(10053):1545–602.
 105. Muldoon MF, Barger SD, Flory JD, Manuck SB. What are quality of life measurements measuring? *Bmj*. 1998;316(7130):542.
 106. Pinto-Prades JL, Attema A, Sánchez-Martínez FI. Measuring Health Utility in Economics. In: *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press; 2019.
 107. Tsiplova K, Pullenayegum E, Cooke T, Xie F. EQ-5D-derived health utilities and minimally important differences for chronic health conditions: 2011 Commonwealth Fund Survey of Sicker Adults in Canada. *Qual Life Res*. 2016 Dec 15;25(12):3009–16.
 108. Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics11Supported by Procter & Gamble Pharmaceuticals, Mason, OH. *Osteoarthr Cartil*. 2004;12:39–44.
 109. MacGregor AJ, Li Q, Spector TD, Williams FMK. The genetic influence on radiographic osteoarthritis is site specific at the hand, hip and knee. *Rheumatology*. 2009;48(3):277–80.
 110. Amanda E. Nelson y Joanne M. Jordan. *Rheumatology*. Seventh Ed. Hochberg, Marc C., MD, MPH, MACP M, editor. Elsevier Inc; 2019. 1503–1512 p.
 111. Segal NA, Felson DT, Torner JC, Zhu Y, Curtis JR, Niu J, et al. Greater trochanteric pain syndrome: epidemiology and associated factors. *Arch Phys Med Rehabil*. 2007 Aug;88(8):988–92.
 112. Khan AM, McLoughlin E, Giannakas K, Hutchinson C, Andrew JG. Hip osteoarthritis: where is the pain? *Ann R Coll Surg Engl*. 2004 Mar;86(2):119–21.
 113. Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B, et al. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum Dis*. 2010 Mar;69(3):483–9.
 114. Kim C, Nevitt MC, Niu J, Clancy MM, Lane NE, Link TM, et al. Association of hip pain with radiographic evidence of hip osteoarthritis: Diagnostic test study. *BMJ*. 2015;351.
 115. 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.
 116. Schiphof D, De Klerk BM, Kerkhof HJM, Hofman A, Koes BW, Boers M, et al. Impact of different descriptions of the Kellgren and Lawrence classification criteria on the diagnosis of knee osteoarthritis. *Ann Rheum Dis*. 2011;70(8):1422–7.
 117. Guermazi A, Hayashi D, Roemer FW, Felson DT. Osteoarthritis. *Rheum Dis Clin North Am*. 2013 Aug;39(3):567–91.
 118. Clayburne G, Baker DG, Schumacher HR. Estimated synovial fluid leukocyte numbers on wet drop preparations as a potential substitute for actual leukocyte counts. *J Rheumatol*. 1992 Jan;19(1):60–2.
 119. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip

- and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis.* 2014;73(7):1323–30.
120. Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan JL, Protheroe J, Jordan KP. Current evidence on risk factors for knee osteoarthritis in older adults: A systematic review and meta-analysis. *Osteoarthr Cartil.* 2015;23(4):507–15.
 121. Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet.* 2016 Oct;388(10053):1545–602.
 122. Nguyen U-SDT, Zhang Y, Zhu Y, Niu J, Zhang B, Felson DT. Increasing prevalence of knee pain and symptomatic knee osteoarthritis: survey and cohort data. *Ann Intern Med.* 2011 Dec 6;155(11):725–32.
 123. Salmon JH, Rat AC, Sellam J, Michel M, Eschard JP, Guillemin F, et al. Economic impact of lower-limb osteoarthritis worldwide: a systematic review of cost-of-illness studies. *Osteoarthr Cartil.* 2016 Sep;24(9):1500–8.
 124. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet.* 2014;384(9945):766–81.
 125. Loza E, Lopez-Gomez JM, Abasolo L, Maese JJ, Carmona L, Batlle-Gualda E. Economic burden of knee and hip osteoarthritis in Spain. *Arthritis Rheum.* 2009;61(2):158–65.
 126. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis.* 2014;73(7):1323–30.
 127. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: Estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis.* 2014;73(7):1323–30.
 128. Clynes MA, Jameson KA, Edwards MH, Cooper C, Dennison EM. Impact of osteoarthritis on activities of daily living: does joint site matter? *Aging Clin Exp Res.* 2019;31(8):1049–56.
 129. Laires PA, Canhã H, Rodrigues AM, Eusébio M, Gouveia M, Branco JC. The impact of osteoarthritis on early exit from work: Results from a population-based study. *BMC Public Health.* 2018;18(1):1–12.
 130. Oliva-Moreno J, Lopez-Bastida J, Worbes-Cerezo M, Serrano-Aguilar P. Health related quality of life of Canary Island citizens. *BMC Public Health.* 2010;10(1):675.
 131. Martín-Fernández J, García -Maroto R, Bilbao A, García-Pérez L, Gutiérrez-Teira B, Molina-Siguero A, et al. Impact of lower limb osteoarthritis on health-related quality of life: A cross-sectional study to estimate the expressed loss of utility in the Spanish population. *PLoS One.* 2020 Jan 24;15(1):e0228398.
 132. Veronese N, Stubbs B, Solmi M, Smith TO, Noale M, Cooper C, et al. Association between lower limb osteoarthritis and incidence of depressive symptoms: data from the osteoarthritis initiative. *Age Ageing.* 2017 May;46(3):470–6.

BIBLIOGRAFÍA

133. Juárez F. El concepto de salud: Una explicación sobre su unicidad, multiplicidad y los modelos de salud. *Int J Psychol Res.* 2011;4(1):70–9.
134. Vega-Franco L. Ideas, creencias y percepciones acerca de la salud. *Reseña histórica. Salud Publica Mex.* 2002;44(3):258–65.
135. Sigerist HE. *Medicine and human welfare.* Yale University Press New Haven; 1941.
136. Svalastog AL, Donev D, Jahren Kristoffersen N, Gajović S. Concepts and definitions of health and health-related values in the knowledge landscapes of the digital society. *Croat Med J.* 2017 Dec;58(6):431–5.
137. World Health Organization. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 1948;
138. Frenk J, Gómez-Dantés O. Designing a framework for the concept of health. *J Public Health Policy.* 2014 Aug 19;35(3):401–6.
139. Gil P. *Medicina preventiva y salud pública.* 10ª. MASSON; 2001. 3,4.
140. Larson JS. The Conceptualization of Health. *Med Care Res Rev.* 1999 Jun 19;56(2):123–36.
141. Dubos R. The state of health and the quality of life. *West J Med.* 1976;125(1):8–9.
142. Terris M. Approaches to an epidemiology of health. *Am J Public Health.* 1975 Oct;65(10):1037–45.
143. Schlaepfer-Pedarzzini L, Infante-Castañeda C, Seblaepfer-Pedrazzini L, Infante-Castañeda C. La medición de salud: perspectivas teóricas y metodológicas. *Salud Publica Mex.* 1990;32(2):141–55.
144. Alonso J. Medidas de salud. *Gestión Sanit Innovaciones y desafíos.* 1997;(CI):2–14.
145. Hutchinson TA, Boyd NF, Feinstein AR. Scientific problems in clinical scales, as demonstrated in the karnofsky index of performance status. *J Chronic Dis.* 1979 Jan;32(9–10):661–6.
146. McHorney CA. Health status assessment methods for adults: Past Accomplishments and Future Challenges 1. *Annu Rev Public Health.* 1999 May;20(1):309–35.
147. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med.* 1993 Apr 15;118(8):622–9.
148. Urzúa M. A. Calidad de vida relacionada con la salud: Elementos conceptuales. *Rev Med Chil.* 2010 Mar;138(3):358–65.
149. Schwartzmann L. Calidad De Vida Relacionada Con La Salud : Health-Related Quality of Life : Conceptual Aspects. *Cienc y Enferm.* 2003;(2):9–21.
150. Karimi M, Brazier J. Health, Health-Related Quality of Life, and Quality of Life: What is the Difference? *Pharmacoeconomics.* 2016 Jul 18;34(7):645–9.
151. Rejeski WJ, Mihalko SL. Physical Activity and Quality of Life in Older Adults. *Journals Gerontol Ser A Biol Sci Med Sci.* 2001 Oct 1;56(Supplement 2):23–35.
152. The WHOQOL Group. The World Health Organization Quality of Life assessment

- (WHOQOL): position paper from the World Health Organization. *Soc Sci Med.* 1995 Nov;41(10):1403–9.
153. Fanshel S, Bush JW. A health-status index and its application to health-services outcomes. *Oper Res.* 1970;18(6):1021–66.
 154. Kaplan RM, Bush JW. Health-related quality of life measurement for evaluation research and policy analysis. *Heal Psychol.* 1982;1(1):61–80.
 155. Hays RD, Anderson R, Revicki D. Psychometric considerations in evaluating health-related quality of life measures. *Qual Life Res.* 1993 Dec;2(6):441–9.
 156. Patrick DL, Erickson P. Health status and health policy: quality of life in health care evaluation and resource allocation. 1993;
 157. Cummins RA. Moving from the quality of life concept to a theory. *J Intellect Disabil Res.* 2005;49(10):699–706.
 158. Ebrahim S. Clinical and public health perspectives and applications of health-related quality of life measurement. *Soc Sci Med.* 1995 Nov;41(10):1383–94.
 159. Andersen RM. National health surveys and the behavioral model of health services use. *Med Care.* 2008;46(7):647–53.
 160. Martín-Fernández J, Ariza-Cardiel G, Polentinos-Castro E, Sanz-Cuesta T, Sarria-Santamera A, del Cura-González I. Explaining differences in perceived health-related quality of life: a study within the Spanish population. *Gac Sanit.* 2018 Sep;32(5):447–53.
 161. Arrospide A, Machón M, Ramos-Goñi JM, Ibarondo O, Mar J. Inequalities in health-related quality of life according to age, gender, educational level, social class, body mass index and chronic diseases using the Spanish value set for Euroqol 5D-5L questionnaire. *Health Qual Life Outcomes.* 2019 Dec 18;17(1):69.
 162. Guyatt GH, Veldhuyzen Van Zanten SJ, Feeny DH, Patrick DL. Measuring quality of life in clinical trials: a taxonomy and review. *CMAJ.* 1989 Jun 15;140(12):1441–8.
 163. Erickson P, Kendall EA, Anderson JP, Kaplan RM. Using Composite Health Status Measures to Assess the Nation's Health. *Med Care.* 1989 Mar;27(Supplement):S66–76.
 164. Wiklund I. The Nottingham Health Profile--a measure of health-related quality of life. *Scand J Prim Health Care Suppl.* 1990;1:15–8.
 165. Parkerson GR, Gehlbach SH, Wagner EH, James SA, Clapp NE, Muhlbaier LH. The Duke-UNC Health Profile: An Adult Health Status Instrument for Primary Care. *Med Care.* 1981 Aug;19(8):806–28.
 166. Nelson E, Wasson J, Kirk J, Keller A, Clark D, Dietrich A, et al. Assessment of function in routine clinical practice: description of the COOP Chart method and preliminary findings. *J Chronic Dis.* 1987;40 Suppl 1:55S-69S.
 167. Donabedian A. The quality of care. How can it be assessed? *JAMA J Am Med Assoc.* 1988 Sep 23;260(12):1743–8.
 168. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy.* 1990 Dec;16(3):199–208.

BIBLIOGRAFÍA

169. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Med Care*. 1992;473–83.
170. Ware JE, Kosinski M, Kellerr SD. A 12-Item Short-Form Health Survey. *Med Care*. 1996 Mar;34(3):220–33.
171. Herdman M, Gudex C, Lloyd a., Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727–36.
172. Badia X, Roset M, Montserrat S, Herdman M, Segura A. [The Spanish version of EuroQol: a description and its applications. European Quality of Life scale]. *Med Clin (Barc)*. 1999;112 Suppl:79–85.
173. Bilbao A, García-Pérez L, Arenaza JC, García I, Ariza-Cardiel G, Trujillo-Martín E, et al. Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. *Qual Life Res*. 2018;27(11):2897–908.
174. Alonso J, Prieto L, Antó JM. [The Spanish version of the SF-36 Health Survey (the SF-36 health questionnaire): an instrument for measuring clinical results]. *Med Clin (Barc)*. 1995 May;104(20):771–6.
175. Vilagut G, Ferrer M, Rajmil L, Rebollo P, Permanyer-Miralda G, Quintana JM, et al. El Cuestionario de Salud SF-36 español: una década de experiencia y nuevos desarrollos. *Gac Sanit*. 2005 Apr;19(2):135–50.
176. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: Results from the IQOLA Project. *J Clin Epidemiol*. 1998;51(11):1171–8.
177. Ware JE, Kosinski M, Dewey JE, Gandek B, others. How to score and interpret single-item health status measures: a manual for users of the SF-8 health survey. Lincoln, RI Qual Inc. 2001;15(10):5.
178. Malet L, Llorca PM, Beringuier B, Lehert P, Falissard B. AIQoL 9 for measuring quality of life in alcohol dependence. *Alcohol Alcohol*. 2006;41(2):181–7.
179. Juniper EF, Sonia Buist A, Cox FM, Ferrie PJ, King DR. Validation of a standardized version of the asthma quality of life questionnaire. *Chest*. 1999;115(5):1265–70.
180. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology. *JNCI J Natl Cancer Inst*. 1993 Mar 3;85(5):365–76.
181. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)—a simple practical measure for routine clinical use. *Clin Exp Dermatol*. 1994 May;19(3):210–6.
182. Bradley C, Todd C, Gorton T, Symonds E, Martin A, Plowright R. The development of an individualized questionnaire measure of perceived impact of diabetes on quality of life: the ADDQoL. *Qual Life Res*. 1999;8(1–2):79–91.
183. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. *Pain*. 1992 Aug;50(2):133–49.
184. Hyland ME, Finnis S, Irvine SH. A scale for assessing quality of life in adult asthma sufferers.

- J Psychosom Res. 1991 Jan;35(1):99–110.
185. van der Molen T, Willemse BWM, Schokker S, ten Hacken NHT, Postma DS, Juniper EF. Development, validity and responsiveness of the clinical COPD questionnaire. *Health Qual Life Outcomes*. 2003;1:1–10.
 186. Badia X, Roca-Cusachs A, Dalfó A, Gascón G, Abellán J, Lahoz R, et al. Validation of the short form of the Spanish Hypertension Quality of Life Questionnaire (MINICHAL). *Clin Ther*. 2002 Dec;24(12):2137–54.
 187. Brodie DA, Inoue A, Shaw DG. Motivational interviewing to change quality of life for people with chronic heart failure: A randomised controlled trial. *Int J Nurs Stud*. 2008 Apr;45(4):489–500.
 188. Badia X, Prieto L, Roset M, Díez-Pérez A. [Development of the ECOS-16 clinical questionnaire for the assessment of the quality of life in patients with osteoporosis]. *Med Clin (Barc)*. 2000;114 Suppl:68–75.
 189. Herdman M, Badia X, Berra S. EuroQol-5D: a simple alternative for measuring health-related quality of life in primary care. *Aten Primaria*. 2001;28(6):425–30.
 190. Ramos-Goñi JM, Craig BM, Oppe M, Ramallo-Fariña Y, Pinto-Prades JL, Luo N, et al. Handling Data Quality Issues to Estimate the Spanish EQ-5D-5L Value Set Using a Hybrid Interval Regression Approach. *Value Health*. 2018;21(5):596–604.
 191. Wailoo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, et al. Mapping to Estimate Health-State Utility from Non-Preference-Based Outcome Measures: An ISPOR Good Practices for Outcomes Research Task Force Report. *Value Heal*. 2017;20(1):18–27.
 192. Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabasés JM, Serrano-Aguilar P, Rivero-Arias O. Valuation and Modeling of EQ-5D-5L Health States Using a Hybrid Approach. *Med Care*. 2017 Jul;55(7):e51–8.
 193. NICE National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. London; 2013.
 194. Dakin H. Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. *Health Qual Life Outcomes*. 2013;11(1):151.
 195. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. *Appl Health Econ Health Policy*. 2017;15(2):127–37.
 196. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002 Mar;21(2):271–92.
 197. Kiadaliri AA, Lamm CJ, de Verdier MG, Engstrom G, Turkiewicz A, Lohmander LS, et al. Association of knee pain and different definitions of knee osteoarthritis with health-related quality of life: a population-based cohort study in southern Sweden. *Health Qual Life Outcomes*. 2016;14(1):121.
 198. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol*. 1988;15(12):1833–40.
 199. Escobar A, Quintana JM, Bilbao A, Azkárate J, Güenaga JI. Validation of the Spanish version

BIBLIOGRAFÍA

- of the WOMAC questionnaire for patients with hip or knee osteoarthritis. Western Ontario and McMaster Universities Osteoarthritis Index. *Clin Rheumatol*. 2002;21(6):466–71.
200. Whitehouse SL, Lingard EA, Katz JN, Learmonth ID. Development and testing of a reduced WOMAC function scale. *J Bone Joint Surg Br*. 2003 Jul;85(5):706–11.
201. Escobar A, Vrotsou K, Bilbao A, Quintana JM, García Pérez L, Herrera-Espiñeira C. Validación de una escala reducida de capacidad funcional del cuestionario WOMAC. *Gac Sanit*. 2011;25(6):513–8.
202. Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee. *Scand J Rheumatol*. 1987;16(SUPPL. 65):85–9.
203. Lequesne MG. The algofunctional indices for hip and knee osteoarthritis. *J Rheumatol*. 1997 Apr;24(4):779–81.
204. Lequesne M. Indices of severity and disease activity for osteoarthritis. *Semin Arthritis Rheum*. 1991 Jun;20(6 Suppl 2):48–54.
205. Basaran S, Guzel R, Seydaoglu G, Guler-Uysal F. Validity, reliability, and comparison of the WOMAC osteoarthritis index and Lequesne algofunctional index in Turkish patients with hip or knee osteoarthritis. *Clin Rheumatol*. 2010 Jul;29(7):749–56.
206. Bae SC, Lee HS, Yun HR, Kim TH, Yoo DH, Kim SY. Cross-cultural adaptation and validation of Korean Western Ontario and McMaster Universities (WOMAC) and Lequesne osteoarthritis indices for clinical research. *Osteoarthr Cartil*. 2001 Nov;9(8):746–50.
207. Xie F, Li S-C, Lo N-N, Yeo S-J, Yang K-Y, Yeo W, et al. Cross-cultural adaptation and validation of Singapore English and Chinese Versions of the Oxford Knee Score (OKS) in knee osteoarthritis patients undergoing total knee replacement. *Osteoarthr Cartil*. 2007 Sep;15(9):1019–24.
208. Stucki G, Sangha O, Stucki S, Michel BA, Tyndall A, Dick W, et al. Comparison of the WOMAC (Western Ontario and McMaster Universities) osteoarthritis index and a self-report format of the self-administered Lequesne-Algofunctional index in patients with knee and hip osteoarthritis. *Osteoarthr Cartil*. 1998 Mar;6(2):79–86.
209. Faucher M, Poiraudeau S, Lefevre-Colau MM, Rannou F, Fermanian J, Revel M. Algo-functional assessment of knee osteoarthritis: comparison of the test-retest reliability and construct validity of the WOMAC and Lequesne indexes. *Osteoarthr Cartil*. 2002 Aug;10(8):602–10.
210. Mahmood T, Choudhury MR, Islam MN, Haq SA, Shahin MA, Ali SMM, et al. Translation, cross-cultural adaptation and validation of the English Lequesne Algofunctional index in to Bengali. *Health Qual Life Outcomes*. 2020;18(1):1–9.
211. Fries JF, Spitz P, Kraines RG, Holman HR. Section of THE Arthritis Foundation Measurement of Patient Outcome in Arthritis. *Arthritis Rheum*. 1980;23(2):137–45.
212. Ramey DR, Raynauld JP, Fries JF. The health assessment questionnaire 1992: status and review. *Arthritis Care Res*. 1992 Sep;5(3):119–29.
213. Batlle E, Chalmeta C. .Calidad de vida en la artritis reumatoide . *Rev Esp Reum*. 2002;1(Supl 1):9-21.
214. Esteve-Vives J, Batlle-Gualda E, Reig A. Spanish version of the Health Assessment

- Questionnaire: reliability, validity and transcultural equivalency. Grupo para la Adaptación del HAQ a la Población Española. *J Rheumatol*. 1993 Dec;20(12):2116–22.
215. Pincus T, Summey JA, Soraci SA, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum*. 1983 Nov;26(11):1346–53.
 216. Wolfe F. Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): analyses in 2,491 rheumatoid arthritis patients following leflunomide initiation. *J Rheumatol*. 2001 May;28(5):982–9.
 217. Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. *Phys Ther*. 1999 Apr;79(4):371–83.
 218. Xu L, Wang C, Zhang C, Feng X, Tong W. Cross-cultural adaption and validation of simplified Chinese version of the lower extremity function scale in patients with knee osteoarthritis. *Clin Rheumatol*. 2020 Oct;39(10):3041–8.
 219. Cruz-Díaz D, Lomas-Vega R, Osuna-Pérez MC, Hita-Contreras F, Fernández ÁD, Martínez-Amat A. The Spanish lower extremity functional scale: A reliable, valid and responsive questionnaire to assess musculoskeletal disorders in the lower extremity. *Disabil Rehabil*. 2014 Nov 5;36(23):2005–11.
 220. Barry MJ, Fowler FJ, O’Leary MP, Bruskewitz RC, Holtgrewe HL, Mebust WK. Measuring disease-specific health status in men with benign prostatic hyperplasia. Measurement Committee of The American Urological Association. *Med Care*. 1995 Apr;33(4 Suppl):AS145-55.
 221. Valentine KD, Brodney S, Vo H, Sepucha K, Fowler FJ, Barry MJ. Validation of the Impact Index: can we measure disease effects on quality of life in patients with hip and knee osteoarthritis? *Qual Life Res*. 2021 Jan 2;
 222. Johanson NA, Liang MH, Daltroy L, Rudicel S, Richmond J. American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments. Reliability, validity, and sensitivity to change. *J Bone Joint Surg Am*. 2004 May;86(5):902–9.
 223. Castellet E, Ares O, Celaya F, Valentí-Azcárate A, Salvador A, Torres A, et al. Transcultural adaptation and validation of the “ Hip and Knee” questionnaire into Spanish. *Health Qual Life Outcomes*. 2014;12(1):1–7.
 224. Rat A-C, Coste J, Pouchot J, Baumann M, Spitz E, Retel-Rude N, et al. OAKHQOL: a new instrument to measure quality of life in knee and hip osteoarthritis. *J Clin Epidemiol*. 2005 Jan;58(1):47–55.
 225. Gonzalez Sáenz De Tejada M, Escobar A, Herdman M, Herrera C, García L, Sarasqueta C. Adaptation and validation of the Osteoarthritis Knee and Hip Quality of Life (OAKHQOL) questionnaire for use in patients with osteoarthritis in Spain. *Clin Rheumatol*. 2011;30(12):1563–75.
 226. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. *J Bone Jt Surg Am*. 1969;51(4):737–55.
 227. Hoeksma HL, Van Den Ende CHM, Runday HK, Heering A, Breedveld FC. Comparison of the responsiveness of the Harris Hip Score with generic measures for hip function in

- osteoarthritis of the hip. *Ann Rheum Dis*. 2003 Oct;62(10):935–8.
228. Söderman P, Malchau H. Is the Harris hip score system useful to study the outcome of total hip replacement? *Clin Orthop Relat Res*. 2001 Mar;(384):189–97.
229. Navarro Collado MJ, Peiró Moreno S, Ruiz Jareño L, Payá Rubio A, Hervás Juan MT, López Matéu P. Validation of the harris hip score in rehabilitation after hip arthroplasty. *Rehabilitacion*. 2005;39(4):147–54.
230. Nilsson AK, Lohmander LS, Klässbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS) - Validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord*. 2003;4:1–8.
231. Goodman SM, Mehta BY, Mandl LA, Szymonifka JD, Finik J, Figgie MP, et al. Validation of the Hip Disability and Osteoarthritis Outcome Score and Knee Injury and Osteoarthritis Outcome Score Pain and Function Subscales for Use in Total Hip Replacement and Total Knee Replacement Clinical Trials. *J Arthroplasty*. 2020 May;35(5):1200-1207.e4.
232. Gandek B, Roos EM, Franklin PD, Ware JE. A 12-item short form of the Hip disability and Osteoarthritis Outcome Score (HOOS-12): tests of reliability, validity and responsiveness. *Osteoarthr Cartil*. 2019;27(5):754–61.
233. Arias-de la Torre J, Puigdomenech E, Valderas JM, Evans JP, Martín V, Molina AJ, et al. Availability of specific tools to assess patient reported outcomes in hip arthroplasty in Spain. Identifying the best candidates to incorporate in an arthroplasty register. A systematic review and standardized assessment. *PLoS One*. 2019;14(4):1–13.
234. Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the hip outcome score. *Arthroscopy*. 2006 Dec;22(12):1304–11.
235. Martin RL, Philippon MJ. Evidence of validity for the hip outcome score in hip arthroscopy. *Arthroscopy*. 2007 Aug;23(8):822–6.
236. Seijas R, Sallent A, Ruiz-Ibán MA, Ares O, Marín-Peña O, Cuéllar R, et al. Validation of the Spanish version of the Hip Outcome Score: a multicenter study. *Health Qual Life Outcomes*. 2014;12:70.
237. Mohtadi NGH, Griffin DR, Pedersen ME, Chan D, Safran MR, Parsons N, et al. The Development and Validation of a Self-Administered Quality-of-Life Outcome Measure for Young, Active Patients With Symptomatic Hip Disease: The International Hip Outcome Tool (iHOT-33). *Arthrosc J Arthrosc Relat Surg*. 2012 May;28(5):595-610.e1.
238. Li DH, Wang W, Li X, Gao YL, Liu DH, Liu DL, et al. Development of a valid Simplified Chinese version of the International Hip Outcome Tool (SC-iHOT-33) in young patients having total hip arthroplasty. *Osteoarthr Cartil*. 2017;25(1):94–8.
239. Tijssen M, Tak I, Stubbe J, Haverkamp D, de Visser E, Nijhuis-van der Sanden M, et al. Translation, Cross-cultural Adaptation, and Validation of the Dutch International Hip Outcome Tool-33 (iHOT-33 NL) in Young, Physically Active Individuals With Symptomatic Hip Joint Pathology. *J Orthop Sports Phys Ther*. 2018 Apr;48(4):289–98.
240. Baumann F, Weber J, Zeman F, Müller M, Lahner M, Nerlich M, et al. Validation of a German version of the International Hip Outcome Tool (G-iHOT33) according to the COSMIN checklist: how much improvement is clinically relevant? *Arch Orthop Trauma Surg*. 2016 Jan;136(1):83–91.

241. Dion M-O, Faure P-A, May O, Bonin N, Beaulé P, Carsen S, et al. Validation of the French version of the self-administered international hip outcome tool-33 questionnaire. *Orthop Traumatol Surg Res.* 2021 Feb;102858.
242. Ruiz-Ibán MA, Seijas R, Sallent A, Ares O, Marín-Peña O, Muriel A, et al. The international Hip Outcome Tool-33 (iHOT-33): Multicenter validation and translation to Spanish. *Health Qual Life Outcomes.* 2015;13(1):1–7.
243. Griffin DR, Parsons N, Mohtadi NGH, Safran MR. A Short Version of the International Hip Outcome Tool (iHOT-12) for Use in Routine Clinical Practice. *Arthrosc J Arthrosc Relat Surg.* 2012 May;28(5):611–8.
244. Stasi S, Stamou M, Papathanasiou G, Frantzeskaki P, Kanavas E, Evaggelou-Sossidis G, et al. International Hip Outcome Tool (12-items) as health-related quality-of-life measure in osteoarthritis: validation of Greek version. *J Patient-Reported Outcomes.* 2020;4(1).
245. Thorborg K, Hölmich P, Christensen R, Petersen J, Roos EM. The Copenhagen Hip and Groin Outcome Score (HAGOS): Development and validation according to the COSMIN checklist. *Br J Sports Med.* 2011;45(6):478–91.
246. Thomeé R, Jónasson P, Thorborg K, Sansone M, Ahldén M, Thomeé C, et al. Cross-cultural adaptation to Swedish and validation of the Copenhagen Hip and Groin Outcome Score (HAGOS) for pain, symptoms and physical function in patients with hip and groin disability due to femoro-acetabular impingement. *Knee Surg Sports Traumatol Arthrosc.* 2014 Apr;22(4):835–42.
247. Brans E, De Graaf JS, Munzebrock AVE, Bessem B, Reininga IHF. Cross-cultural adaptation and validation of the Dutch version of the hip and groin outcome score (HAGOS-NL). *PLoS One.* 2016;11(1):1–12.
248. Cao S, Cao J, Li S, Wang W, Qian Q, Ding Y. Cross-cultural adaptation and validation of the Simplified Chinese version of Copenhagen Hip and Groin Outcome Score (HAGOS) for total hip arthroplasty. *J Orthop Surg Res.* 2018 Nov 6;13(1):278.
249. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Jt Surg.* 1996 Jan 1;78-B(2):185–90.
250. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr a J, et al. The use of the Oxford hip and knee scores. *J Bone Jt Surg.* 2007;89(8):1010–4.
251. Dawson J, Fitzpatrick R, Murray D, Carr A. Comparison of measures to assess outcomes in total hip replacement surgery. *Qual Saf Heal Care.* 1996 Jun 1;5(2):81–8.
252. Kalirajah Y, Azurza K, Hulme C, Molloy S, Drabu KJ. Health outcome measures in the evaluation of total hip arthroplasties - A comparison between the harris hip score and the Oxford hip score. *J Arthroplasty.* 2005;20(8):1037–41.
253. Gosens T, Hoefnagels NHM, de Vet RCW, Dhert WJA, van Langelaan EJ, Bulstra SK, et al. The "Oxford Heup Score." *Acta Orthop.* 2005;76(2):204–11.
254. Delaunay C, Epinette JA, Dawson J, Murray D, Jolles BM. Cross-cultural adaptations of the Oxford-12 HIP score to the French speaking population. *Orthop Traumatol Surg Res.* 2009;95(2):89–99.
255. Naal FD, Sieverding M, Impellizzeri FM, von Knoch F, Mannion AF, Leunig M. Reliability and

BIBLIOGRAFÍA

- Validity of the Cross-Culturally Adapted German Oxford Hip Score. *Clin Orthop Relat Res.* 2009 Apr 26;467(4):952–7.
256. Martinelli N, Longo UG, Marinozzi A, Franceschetti E, Costa V, Denaro V. Cross-cultural adaptation and validation with reliability, validity, and responsiveness of the Italian version of the Oxford Hip Score in patients with hip osteoarthritis. *Qual Life Res.* 2011 Aug 7;20(6):923–9.
257. Paulsen A, Odgaard A, Overgaard S. Translation, cross-cultural adaptation and validation of the Danish version of the Oxford hip score: Assessed against generic and disease-specific questionnaires. *Bone Jt Res.* 2012 Sep 19;1(9):225–33.
258. Tuğay BU, Tuğay N, Güney H, Hazar Z, Yüksel İ, Atilla B. Cross-cultural adaptation and validation of the Turkish version of Oxford hip score. *Arch Orthop Trauma Surg.* 2015 Jun 11;135(6):879–89.
259. Lee Y-K, Chung CY, Park MS, Lee KM, Koo K-H, Lee DJ, et al. Transcultural adaptation and testing of psychometric properties of the Korean version of the Oxford hip score. *J Orthop Sci.* 2012 Jul;17(4):377–81.
260. Uesugi Y, Makimoto K, Fujita K, Nishii T, Sakai T, Sugano N. Validity and responsiveness of the Oxford hip score in a prospective study with Japanese total hip arthroplasty patients. *J Orthop Sci.* 2009;14(1):35–9.
261. Haragus H, Prejbeanu R, Poenaru D V., Deleanu B, Timar B, Vermesan D. Cross-cultural adaptation and validation of a patient-reported hip outcome score. *Int Orthop.* 2018 May 6;42(5):1001–6.
262. Al-Samhan DA, Al-Mutairi DO, Al-Kudair DA, Khaja DA. Translation and Cross-Cultural Adaptation of the Oxford Hip Score into Arabic. *Glob J Med Res.* 2020 May 9;(April):19–26.
263. Martínez JP, Arango AS, Castro AM, Martínez Rondanelli A. Validación de la versión en español de las escalas de Oxford para rodilla y cadera. *Rev Colomb Ortop y Traumatol.* 2016;30(2):61–6.
264. Hawker GA, Davis AM, French MR, Cibere J, Jordan JM, March L, et al. Development and preliminary psychometric testing of a new OA pain measure--an OARSI/OMERACT initiative. *Osteoarthr Cartil.* 2008 Apr;16(4):409–14.
265. Erel S, Şimşek İE, Özkan H. Analysis of the reliability and validity of the Turkish version of the intermittent and constant osteoarthritis pain questionnaire. *Acta Orthop Traumatol Turc.* 2015;49(5):508–12.
266. Gonçalves RS, Cabri J, Pinheiro JP, Ferreira PL, Gil J. Cross-cultural adaptation and validation of the Portuguese version of the intermittent and constant osteoarthritis pain (ICOAP) measure for the knee. *Osteoarthr Cartil.* 2010 Aug;18(8):1058–61.
267. Kessler S, Grammosis A, Günther K-P, Kirschner S. [The intermittent and constant pain score (ICOAP) - a questionnaire to assess pain in patients with gonarthrosis]. *Z Orthop Unfall.* 2011 Jan;149(1):22–6.
268. Manolarakis GE, Kontodimopoulos N, Sifaki-Pistolla D, Niakas D. Establishing the Psychometric Properties of the ICOAP Questionnaire through Intra-Articular Treatment of Osteoarthritic Pain: Implementation for the Greek Version. *Arthritis.* 2016;2016:6201802.
269. Zhang C, Liu DH, Qu YL, Jia ZY, Wang W, Li J, et al. Transcultural adaptation and validation

- of the Chinese version of the intermittent and constant osteoarthritis pain (ICOAP) measure in patients with knee osteoarthritis. *Osteoarthr Cartil.* 2017;25(4):506–12.
270. Hefti F, Müller W, Jakob RP, Stäubli HU. Evaluation of knee ligament injuries with the IKDC form. *Knee Surg Sports Traumatol Arthrosc.* 1993;1(3–4):226–34.
271. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, Neyret P, et al. Development and validation of the international knee documentation committee subjective knee form. *Am J Sports Med.* 2001;29(5):600–13.
272. Padua R, Bondi R, Ceccarelli E, Bondi L, Romanini E, Zanoli G, et al. Italian version of the international knee documentation committee subjective knee form: Cross-cultural adaptation and validation. *Arthrosc J Arthrosc Relat Surg.* 2004 Oct;20(8):819–23.
273. Koumantakis GA, Tsofigkas K, Papoutsidakis A, Ververidis A, Drosos GI. Cross-cultural adaptation and validation of the International Knee Documentation Committee Subjective Knee Form in Greek. *J Orthop Traumatol.* 2016 Jun;17(2):123–9.
274. Haverkamp D, Sierevelt IN, Breugem SJM, Lohuis K, Blankevoort L, van Dijk CN. Translation and validation of the Dutch version of the International Knee Documentation Committee Subjective Knee Form. *Am J Sports Med.* 2006 Oct;34(10):1680–4.
275. Lertwanich P, Praphruetkit T, Keyurapan E, Lamsam C, Kulthanan T. Validity and reliability of Thai version of the International Knee Documentation Committee Subjective Knee Form. *J Med Assoc Thai.* 2008 Aug;91(8):1218–25.
276. Metsavaht L, Leporace G, Riberto M, de Mello Sposito MM, Batista LA. Translation and cross-cultural adaptation of the Brazilian version of the International Knee Documentation Committee Subjective Knee Form: validity and reproducibility. *Am J Sports Med.* 2010 Sep;38(9):1894–9.
277. Fu S-N, Chan Y-H. Translation and validation of Chinese version of International Knee Documentation Committee Subjective Knee Form. *Disabil Rehabil.* 2011;33(13–14):1186–9.
278. Kim JG, Ha JK, Lee JY, Seo SS, Choi CH, Lee MC. Translation and validation of the Korean version of the international knee documentation committee subjective knee form. *Knee Surg Relat Res.* 2013 Sep;25(3):106–11.
279. Ebrahimzadeh MH, Makhmalbaf H, Golhasani-Keshtan F, Rabani S, Birjandinejad A. The International Knee Documentation Committee (IKDC) Subjective Short Form: a validity and reliability study. *Knee Surg Sports Traumatol Arthrosc.* 2015 Nov;23(11):3163–7.
280. Çelik D, Coşkunsu D, Kiliçoğlu Ö, Ergönül Ö, Irrgang JJ. Translation and cross-cultural adaptation of the international knee documentation committee subjective knee form into Turkish. *J Orthop Sports Phys Ther.* 2014 Nov;44(11):899–909.
281. Huang H, Nagao M, Arita H, Nishio H, Kaneko H, Saita Y, et al. Validation and defining the minimal clinically important difference of the Japanese version of the IKDC Subjective Knee Form. *J Orthop Sci.* 2021 Jan;26(1):149–55.
282. 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 Sport Phys Ther.* 1998 Aug;28(2):88–96.

BIBLIOGRAFÍA

283. Ornetti P, Parratte S, Gossec L, Tavernier C, Argenson J-N, Roos EM, et al. Cross-cultural adaptation and validation of the French version of the Knee injury and Osteoarthritis Outcome Score (KOOS) in knee osteoarthritis patients. *Osteoarthr Cartil.* 2008 Apr;16(4):423–8.
284. Nakamura N, Takeuchi R, Sawaguchi T, Ishikawa H, Saito T, Goldhahn S. Cross-cultural adaptation and validation of the Japanese Knee Injury and Osteoarthritis Outcome Score (KOOS). *J Orthop Sci.* 2011 Sep;16(5):516–23.
285. Gonçalves RS, Cabri J, Pinheiro JP, Ferreira PL. Cross-cultural adaptation and validation of the Portuguese version of the Knee injury and Osteoarthritis Outcome Score (KOOS). *Osteoarthr Cartil.* 2009 Sep;17(9):1156–62.
286. Salavati M, Mazaheri M, Negahban H, Sohani SM, Ebrahimian MR, Ebrahimi I, et al. Validation of a Persian-version of Knee injury and Osteoarthritis Outcome Score (KOOS) in Iranians with knee injuries. *Osteoarthr Cartil.* 2008 Oct;16(10):1178–82.
287. de Groot IB, Favejee MM, Reijman M, Verhaar JAN, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. *Health Qual Life Outcomes.* 2008 Feb 26;6:16.
288. Xie F, Li S-C, Roos EM, Fong K-Y, Lo N-N, Yeo S-J, et al. Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Knee injury and Osteoarthritis Outcome Score (KOOS) in Asians with knee osteoarthritis in Singapore. *Osteoarthr Cartil.* 2006 Nov;14(11):1098–103.
289. 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 Surgery, Sport Traumatol Arthrosc.* 2014 Jan 10;22(1):104–8.
290. Perruccio A V., Stefan Lohmander L, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS) - an OARSI/OMERACT initiative. *Osteoarthr Cartil.* 2008;16(5):542–50.
291. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res.* 1989 Nov;(248):13–4.
292. Scuderi GR, Bourne RB, Noble PC, Benjamin JB, Lonner JH, Scott WN. The new knee society knee scoring system. *Clin Orthop Relat Res.* 2012;470(1):3–19.
293. Ghanem E, Pawasarat I, Lindsay A, May L, Azzam K, Joshi A, et al. Limitations of the Knee Society Score in Evaluating Outcomes Following Revision Total Knee Arthroplasty. *J Bone Jt Surgery-American Vol.* 2010 Oct;92(14):2445–51.
294. Ares O, León-Muñoz VJ, Castellet E, Pelfort X, Hinarejos P, Amillo JR, et al. Translation and validation of the new knee society knee scoring system into Spanish: Spanish KSS translation. *Knee Surgery, Sport Traumatol Arthrosc.* 2020 Oct 8;
295. Scuderi GR, Sikorskii A, Bourne RB, Lonner JH, Benjamin JB, Noble PC. The Knee Society Short Form Reduces Respondent Burden in the Assessment of Patient-reported Outcomes. *Clin Orthop Relat Res.* 2016;474(1):134–42.
296. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients

- about total knee replacement. *J Bone Joint Surg Br.* 1998;80(1):63–9.
297. Garratt AM, Brealey S, Gillespie WJ. Patient-assessed health instruments for the knee: A structured review. *Rheumatology.* 2004;43(11):1414–23.
 298. Benson T, Williams DH, Potts HWW. Performance of EQ-5D, howRu and Oxford hip & knee scores in assessing the outcome of hip and knee replacements. *BMC Health Serv Res.* 2016 Dec 22;16(1):512.
 299. Goldhahn S, Takeuchi R, Nakamura N, Nakamura R, Sawaguchi T. Responsiveness of the Knee Injury and Osteoarthritis Outcome Score (KOOS) and the Oxford Knee Score (OKS) in Japanese patients with high tibial osteotomy. *J Orthop Sci.* 2017 Sep;22(5):862–7.
 300. Padua R, Zanolli G, Ceccarelli E, Romanini E, Bondi R, Campi A. The Italian version of the Oxford 12-item Knee Questionnaire?cross-cultural adaptation and validation. *Int Orthop.* 2003 Aug 1;27(4):214–6.
 301. Haverkamp D, Breugem SJM, Sierevelt IN, Blankevoort L, van Dijk CN. Translation and validation of the Dutch version of the Oxford 12-item knee questionnaire for knee arthroplasty. *Acta Orthop.* 2005;76(3):347–52.
 302. Xie F, Thumboo J, Lo N-N, Yeo S-J, Yang K-Y, Yeo W, et al. Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Lequesne Algofunctional Index of knee in Asians with knee osteoarthritis in Singapore. *Osteoarthr Cartil.* 2007 Jan;15(1):19–26.
 303. Naal FD, Impellizzeri FM, Sieverding M, Loibl M, von Knoch F, Mannion AF, et al. The 12-item Oxford Knee Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with osteoarthritis of the knee. *Osteoarthr Cartil.* 2009;17(1):49–52.
 304. Jenny JY, Diesinger Y. Validation of a French version of the Oxford knee questionnaire. *Orthop Traumatol Surg Res.* 2011;97(3):267–71.
 305. Takeuchi R, Sawaguchi T, Nakamura N, Ishikawa H, Saito T, Goldhahn S. Cross-cultural adaptation and validation of the Oxford 12-item knee score in Japanese. *Arch Orthop Trauma Surg.* 2011 Feb 10;131(2):247–54.
 306. Gonçalves RS, Tomás AM, Martins DI. Cross-cultural adaptation and validation of the Portuguese version of the Oxford Knee Score (OKS). *Knee.* 2012;19(4):344–7.
 307. Eun IS, Kim OG, Kim CK, Lee HS, Lee JS. Validation of the Korean Version of the Oxford Knee Score in Patients Undergoing Total Knee Arthroplasty. *Clin Orthop Relat Res.* 2013 Feb 11;471(2):600–5.
 308. Ebrahimzadeh MH, Makhmalbaf H, Birjandinejad A, Soltani-Moghaddas SH. Cross-cultural adaptation and validation of the persian version of the oxford knee score in patients with knee osteoarthritis. *Iran J Med Sci.* 2014;39(6):529–35.
 309. Strimpakos N, Dapka F, Papachristou A, Kapreli E. The 12-item oxford knee score: cross-cultural adaptation into Greek and assessment of its psychometric properties. *Physiotherapy.* 2015;101:e1445–6.
 310. Alghadir AH, Al-Eisa ES, Anwer S. Cross-cultural adaptation and psychometric analysis of the Arabic version of the oxford knee score in adult male with knee osteoarthritis. *BMC*

BIBLIOGRAFÍA

- Musculoskelet Disord. 2017;18(1):190.
311. Reito A, Järivistö A, Jämsen E, Skyttä E, Remes V, Huhtala H, et al. Translation and validation of the 12-item Oxford knee score for use in Finland. *BMC Musculoskelet Disord.* 2017;18(1):74.
 312. Prieto L, Badia X. Cuestionarios de salud : concepto y metodología. *Atención Primaria.* 2001;28(3):201–9.
 313. Oosterveld P, Vorst HCM, Smits N. Methods for questionnaire design: a taxonomy linking procedures to test goals. *Qual Life Res.* 2019 Sep 18;28(9):2501–12.
 314. Carvajal A, Centeno C, Watson R, Martínez M, Sanz Rubiales Á. ¿Cómo validar un instrumento de medida de la salud? *An Sist Sanit Navar.* 2011 Apr;34(1):63–72.
 315. Arafat S, Chowdhury H, Qusar M, Hafez M. Cross Cultural Adaptation and Psychometric Validation of Research Instruments: a Methodological Review. *J Behav Heal.* 2016;5(3):129.
 316. McDowell I. *Measuring health: A guide to rating scales and questionnaires.* 3th ed. Oxford University Press,; 2006.
 317. Novick MR. The axioms and principal results of classical test theory. *J Math Psychol.* 1966;3(1):1–18.
 318. Hambleton RK, Van Der Linden WJ. *Advances in Item Response Theory and Applications: An Introduction.* *Appl Psychol Meas.* 1982;6(4):373–8.
 319. DeVellis RF. *Classical Test Theory.* *Med Care.* 2006 Nov 15;44(Suppl 3):S50–9.
 320. Rasch G. Probabilistic models for some intelligence and attainment tests. ERIC; 1993.
 321. Petrillo J, Cano SJ, McLeod LD, Coon CD. Using Classical Test Theory, Item Response Theory, and Rasch Measurement Theory to Evaluate Patient-Reported Outcome Measures: A Comparison of Worked Examples. *Value Heal.* 2015 Jan;18(1):25–34.
 322. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737–45.
 323. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, 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.
 324. Šimkovic M, Träuble B. Robustness of statistical methods when measure is affected by ceiling and/or floor effect. Hutson AD, editor. *PLoS One.* 2019 Aug 19;14(8):e0220889.
 325. De Boer MR, De Vet HCW, Terwee CB, Moll AC, Völker-Dieben HJM, Van Rens GHMB. Changes to the subscales of two vision-related quality of life questionnaires are proposed. *J Clin Epidemiol.* 2005;58(12):1260–8.
 326. Argimón Pallás JM, Jiménez Villa J. Validación De Cuestionarios. In: Argimón Pallás JM, Jiménez Villa J, editors. *Métodos de investigación clínica y epidemiológica.* 5ª. Elsevier; 2019. p. 231–43.
 327. Palazzo C, Nguyen C, Lefevre-Colau MM, Rannou F, Poiraudau S. Risk factors and burden

- of osteoarthritis. *Ann Phys Rehabil Med*. 2016;59(3):134–8.
328. Prieto-Alhambra D, Judge A, Javaid MK, Cooper C, Diez-Perez A, Arden NK. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: Influences of age, gender and osteoarthritis affecting other joints. *Ann Rheum Dis*. 2014;73(9):1659–64.
329. Wallace IJ, Worthington S, Felson DT, Jurmain RD, Wren KT, Maijanen H, et al. Knee osteoarthritis has doubled in prevalence since the mid-20th century. *Proc Natl Acad Sci*. 2017 Aug 29;114(35):9332–6.
330. Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, et al. OARSI recommendations for the management of hip and knee osteoarthritis. Part III: Changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthr Cartil*. 2010;18(4):476–99.
331. Altman RD. Criteria for classification of clinical osteoarthritis. *J Rheumatol Suppl*. 1991 Feb;27:10–2.
332. Green SB. How Many Subjects Does It Take To Do A Regression Analysis. *Multivariate Behav Res*. 1991 Jul;26(3):499–510.
333. Wolf EJ, Harrington KM, Clark SL, Miller MW. Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educ Psychol Meas*. 2013 Dec 1;73(6):913–34.
334. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med*. 2002;21(9):1331–5.
335. Instituto Nacional de Estadística. Ministerio de Sanidad Servicios Sociales e Igualdad. Metodología de la Encuesta Nacional de Salud 2011–2012. 2013. p. 1–12.
336. Bilbao A, García-Pérez L, Arenaza JC, García I, Ariza-Cardiel G, Trujillo-Martín E, et al. Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. *Qual Life Res*. 2018 Nov;27(11):2897–908.
337. Luque R, Rizo B, Urda A, Garcia-Crespo R, Moro E, Marco F, et al. Predictive factors for failure after total knee replacement revision. *Int Orthop*. 2014;38(2):429–35.
338. Abdel Messih M, Naylor JM, Descallar J, Manickam A, Mittal R, Harris IA. Mail Versus Telephone Administration of the Oxford Knee and Hip Scores. *J Arthroplasty*. 2014 Mar;29(3):491–4.
339. Domingo-Salvany A, Bacigalupe A, Carrasco JM, Espelt A, Ferrando J, Borrell C. Propuestas de clase social neoweberiana y neomarxista a partir de la Clasificación Nacional de Ocupaciones 2011. *Gac Sanit*. 2013 May;27(3):263–72.
340. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373–83.
341. Charlson ME, Charlson RE, Peterson JC, Marinopoulos SS, Briggs WM, Hollenberg JP. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. *J Clin Epidemiol*. 2008;61(12):1234–40.
342. Escobar A, Quintana JM, Bilbao A, Azkárate J, Güenaga JI. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. *Western Ontario*

BIBLIOGRAFÍA

- and McMaster Universities Osteoarthritis Index. *Clin Rheumatol.* 2002;21(6):466–71.
343. Jones AM. *Models for health care.* University of York., Centre for Health Economics; 2010.
344. Hardin J, Hilbe J. *Generalized Linear Models and Extensions.* Third Edit. College Station: Stata Press; 2012.
345. Blough DK, Madden CW, Hornbrook MC. Modeling risk using generalized linear models. *J Health Econ.* 1999;18(2):153–71.
346. Long JS, Ervin LH. Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *Am Stat.* 2000;54(3):217–24.
347. Fabozzi FJ, Focardi SM, Rachev ST, Arshanapalli BG. Appendix E: Model Selection Criterion: AIC and BIC. In: *The Basics of Financial Econometrics.* Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2014. p. 399–403.
348. Kass RE, Raftery AE. Bayes Factors. *J Am Stat Assoc.* 1995 Jun;90(430):773–95.
349. Beavers AS, Lounsbury JW, Richards JK, Huck SW, Skolits GJ, Esquivel SL. Practical considerations for using exploratory factor analysis in educational research. *Pract Assessment, Res Eval.* 2013;18(6):1–13.
350. Henson RK, Roberts JK. Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice. *Educ Psychol Meas.* 2006;66(3):393–416.
351. Batista-Foguet JM, Coenders G, Alonso J. Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud. *Med Clin (Barc).* 2004;122(Supl.1):21–7.
352. Mulaik SA, James LR, Van Alstine J, Bennett N, Lind S, Stilwell CD. Evaluation of goodness-of-fit indices for structural equation models. *Psychol Bull.* 1989;105(3):430.
353. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review. *J Educ Res.* 2006;99(6):323–38.
354. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951 Sep;16(3):297–334.
355. Cortina JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol.* 1993;78(1):98–104.
356. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics.* 1977;33(2):363–74.
357. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care.* 1989;27(3):S178--S189.
358. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol.* 2003;56(5):395–407.
359. Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol.* 2004;57(10):1008–18.
360. Weir JP. The intraclass Correlation Coefficient and the SEM. *J Strength Cond Res.*

2005;19(1):231–40.

361. Ministerio de Sanidad Consumo y Bienestar social. Encuesta Nacional de Salud ENSE, España 2017 Serie informes monográficos #1-SALUD MENTAL. 2017.
362. Bernal-Delgado E, García-Armesto S, Oliva J, Sánchez-Martínez F-I, Repullo JR, Peña-Longobardo LM, et al. Spain Health System Review. *Health Syst Transit*. 2018;20(2):1–179.
363. Juhakoski R, Heliövaara M, Impivaara O, Kröger H, Knekt P, Lauren H, et al. Risk factors for the development of hip osteoarthritis: a population-based prospective study. *Rheumatology (Oxford)*. 2009 Jan;48(1):83–7.
364. Wu M, Brazier JE, Kearns B, Relton C, Smith C, Cooper CL. Examining the impact of 11 long-standing health conditions on health-related quality of life using the EQ-5D in a general population sample. *Eur J Heal Econ*. 2015 Mar 10;16(2):141–51.
365. Salaffi F, Di Carlo M, Carotti M, Farah S, Ciapetti A, Gutierrez M. The impact of different rheumatic diseases on health-related quality of life: a comparison with a selected sample of healthy individuals using SF-36 questionnaire, EQ-5D and SF-6D utility values. *Acta Biomed*. 2019 Jan 15;89(4):541–57.
366. Bernad-Pineda M, de las Heras-Sotos J, Garcés-Puentes M V. Calidad de vida en pacientes con artrosis de rodilla y/o cadera. *Rev Esp Cir Ortop Traumatol*. 2014;58(5):283–9.
367. van Schoor NM, Zambon S, Castell M V., Cooper C, Denkiner M, Dennison EM, et al. Impact of clinical osteoarthritis of the hip, knee and hand on self-rated health in six European countries: the European Project on OSteoArthritis. *Qual Life Res*. 2016 Jun 7;25(6):1423–32.
368. Carlesso LC, Hawker GA, Waugh EJ, Davis AM. Disease-specific pain and function predict future pain impact in hip and knee osteoarthritis. *Clin Rheumatol*. 2016 Dec 6;35(12):2999–3005.
369. Serra-Sutton V, Allepuz A, Martínez O, Espallargues M. Factores relacionados con la calidad de vida al año de la artroplastia total de cadera y rodilla: Estudio multicéntrico en Cataluña. *Rev Esp Cir Ortop Traumatol*. 2013;57(4):254–62.
370. Giron P. Determinants of self-rated health in Spain: differences by age groups for adults. *Eur J Public Health*. 2012 Feb 1;22(1):36–40.
371. Otero-Rodríguez A, León-Muñoz LM, Banegas JR, Guallar-Castillón P, Rodríguez-Artalejo F, Regidor E. Life-course socioeconomic position and change in quality of life among older adults: evidence for the role of a critical period, accumulation of exposure and social mobility. *J Epidemiol Community Health*. 2011;65(11):964–71.
372. Sprangers M a, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model. *Soc Sci Med*. 1999;48(11):1507–15.
373. Wilkie R, Blagojevic-Bucknall M, Jordan KP, Pransky G. Onset of work restriction in employed adults with lower limb joint pain: Individual factors and area-level socioeconomic conditions. *J Occup Rehabil*. 2013;23(2):180–8.
374. Castaño Carou A, Pita Fernández S, Pértega Díaz S, de Toro Santos FJ. Clinical Profile, Level of Affection and Therapeutic Management of Patients With Osteoarthritis in Primary Care: The Spanish Multicenter Study EVALÚA. *Reumatol Clínica (English Ed)*. 2015 Nov;11(6):353–

BIBLIOGRAFÍA

- 60.
375. García-Pérez L, Ramos-García V, Serrano-Aguilar P, Pais-Brito JL, Aciego de Mendoza M, Martín-Fernández J, et al. EQ-5D-5L utilities per health states in Spanish population with knee or hip osteoarthritis. *Health Qual Life Outcomes*. 2019 Dec 30;17(1):164.
376. Ostendorf M, Van Stel HF, Buskens E, Schrijvers a. JP, Marting LN, Verbout a. J, et al. Patient-reported outcome in total hip replacement. *J Bone Jt Surg*. 2004;86(6):801–8.
377. Norquist JM, Fitzpatrick R, Dawson J, Jenkinson C. Comparing alternative Rasch-based methods vs raw scores in measuring change in health. *Med Care*. 2004;42(1):1–25.
378. Harris KK, Price AJ, Beard DJ, Fitzpatrick R, Jenkinson C, Dawson J. Can pain and function be distinguished in the Oxford Hip Score in a meaningful way? : an exploratory and confirmatory factor analysis. *Bone Joint Res*. 2014;3(11):305–9.
379. Dawson J, Fitzpatrick R, Frost S, Gundle R, McLardy-Smith P, Murray D. Evidence for the validity of a patient-based instrument for assessment of outcome after revision hip replacement. *J Bone Joint Surg Br*. 2001;83:1125–9.
380. Beard DJ, Harris K, Dawson J, Doll H, Murray DW, Carr AJ, et al. Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. *J Clin Epidemiol*. 2015;68(1):73–9.
381. Norman GR, Sloan J a, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41(5):582–92.
382. Strand LI, Ljunggren AE, Bogen B, Ask T, Johnsen TB. The Short-Form McGill Pain Questionnaire as an outcome measure: Test-retest reliability and responsiveness to change. *Eur J Pain*. 2008;12(7):917–25.
383. Weel H, Lindeboom R, Kuipers SE, Vervest TMJS. Comparison between the Harris- and Oxford Hip Score to evaluate outcomes one-year after total hip arthroplasty. *Acta Orthop Belg*. 2017;83(1):98–109.
384. Yeo MGH, Goh GS, Chen JY, Lo N-N, Yeo S-J, Liow MHL. Are Oxford Hip Score and Western Ontario and McMaster Universities Osteoarthritis Index Useful Predictors of Clinical Meaningful Improvement and Satisfaction After Total Hip Arthroplasty? *J Arthroplasty*. 2020 Sep;35(9):2458–64.
385. Hamilton DF, Lane J V, Gaston P, Patton JT, Macdonald D, Simpson AHRW, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open*. 2013;3(4).
386. Barrack RL, Ruh EL, Chen J, Lombardi A V, Berend KR, Parvizi J, et al. Impact of socioeconomic factors on outcome of total knee arthroplasty. *Clin Orthop Relat Res*. 2014 Jan;472(1):86–97.
387. Hamilton DF, Loth FL, MacDonald DJ, Giesinger K, Patton JT, Simpson AH, et al. Treatment Success Following Joint Arthroplasty: Defining Thresholds for the Oxford Hip and Knee Scores. *J Arthroplasty*. 2018 Aug;33(8):2392–7.
388. Thorborg K, Roos E, Bartels E, Petersen J, Holmich P. Validity, reliability and responsiveness of patient-reported outcome questionnaires when assessing hip and groin disability: a systematic review. *Br J Sports Med*. 2010 Dec 1;44(16):1186–96.

389. Impellizzeri FM, Mannion AF, Naal FD, Leunig M. Validity, reproducibility, and responsiveness of the oxford hip score in patients undergoing surgery for femoroacetabular impingement. *Arthroscopy*. 2015;31(1):42–50.
390. Ares O, Castellet E, Maculé F, León V, Montañez E, Freire A, et al. Translation and validation of “The Knee Society Clinical Rating System” into Spanish. *Knee Surgery, Sport Traumatol Arthrosc*. 2013;21(11):2618–24.
391. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol*. 1999;52(9):861–73.
392. Collins NJ, Prinsen CAC, Christensen R, Bartels EM, Terwee CB, Roos EM. Knee Injury and Osteoarthritis Outcome Score (KOOS): Systematic review and meta-analysis of measurement properties. *Osteoarthr Cartil*. 2015;24(8):1317–29.
393. Harris K, Dawson J, Doll H, Field RE, Murray DW, Fitzpatrick R, et al. Can pain and function be distinguished in the Oxford Knee Score in a meaningful way? An exploratory and confirmatory factor analysis. *Qual Life Res*. 2013;22(9):2561–8.
394. Gooberman-Hill R, Woolhead G, MacKichan F, Ayis S, Williams S, Dieppe P. Assessing chronic joint pain: Lessons from a focus group study. *Arthritis Rheum*. 2007 May 15;57(4):666–71.
395. Heuts PHTG, Vlaeyen JWS, Roelofs J, de Bie RA, Aretz K, van Weel C, et al. Pain-related fear and daily functioning in patients with osteoarthritis. *Pain*. 2004 Jul;110(1–2):228–35.
396. Boersma K, Linton SJ. How does persistent pain develop? An analysis of the relationship between psychological variables, pain and function across stages of chronicity. *Behav Res Ther*. 2005 Nov;43(11):1495–507.
397. Pinedo-Villanueva R, Khalid S, Wylde V, Gooberman-Hill R, Soni A, Judge A. Identifying individuals with chronic pain after knee replacement: a population-cohort, cluster-analysis of Oxford knee scores in 128,145 patients from the English National Health Service. *BMC Musculoskelet Disord*. 2018 Dec 2;19(1):354.
398. Goldhahn S, Takeuchi R, Nakamura N, Nakamura R, Sawaguchi T. Responsiveness of the Knee Injury and Osteoarthritis Outcome Score (KOOS) and the Oxford Knee Score (OKS) in Japanese patients with high tibial osteotomy. *J Orthop Sci*. 2017 Jun;
399. Clement ND, MacDonald D, Simpson A. The minimal clinically important difference in the Oxford knee score and Short Form 12 score after total knee arthroplasty. *Knee Surgery, Sport Traumatol Arthrosc*. 2014;22(8):1933–9.
400. Clement ND, Jenkins PJ, Macdonald D, Nie YX, Patton JT, Breusch SJ, et al. Socioeconomic status affects the Oxford knee score and Short-Form 12 score following total knee replacement. *Bone Joint J*. 2013;95-B(1):52–8.
401. Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. *BMJ Open*. 2012;2(1):e000435.

BIBLIOGRAFÍA

10. Anexos

Anexo 1. Versión adaptada al español (España) del Oxford Hip Score.

1. Durante las últimas 4 semanas, ¿cómo describiría el dolor que tiene normalmente debido a la cadera?

Ninguno	Muy leve	Leve	Moderado	Grave
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. Durante las últimas 4 semanas ¿ha tenido algún problema para lavarse y secarse (todo el cuerpo) debido a su cadera?

Ningún problema	Muy pocos problemas	Problemas moderados	Dificultad extrema	Imposible hacerlo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3. Durante las últimas 4 semanas, ¿ha tenido algún problema para subir y bajar de un coche o para usar el transporte público debido a su cadera? (sea cual sea el que suele utilizar)

Ningún problema	Muy pocos problemas	Problemas moderados	Dificultad extrema	Imposible hacerlo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4. Durante las últimas 4 semanas, ¿se ha podido poner un par de calcetines, medias o pantis?

Sí, fácilmente	Con poca dificultad	Con dificultad moderada	Con extrema dificultad	No, imposible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5. Durante las últimas 4 semanas, ¿podría hacer la compra para su casa usted solo/a?

Sí, fácilmente	Con poca dificultad	Con dificultad moderada	Con extrema dificultad	No, imposible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6. Durante las últimas 4 semanas, ¿durante cuánto tiempo ha podido andar antes de que el dolor debido a su cadera se volviera grave (con o sin bastón)?

Sin dolor/Más de 30 minutos	16 a 30 minutos	5 a 15 minutos	Solo por casa	Nada/dolor grave al caminar
-----------------------------	-----------------	----------------	---------------	-----------------------------

ANEXOS

7. Durante las últimas 4 semanas, ¿ha podido subir un tramo de escaleras?

Sí, fácilmente Con poca dificultad Con dificultad moderada Con extrema dificultad No, imposible

8. Durante las últimas 4 semanas, tras una comida (sentado/a a una mesa), ¿en qué medida le ha resultado doloroso levantarse de la silla debido a su cadera?

Nada doloroso Ligeramente doloroso Moderadamente doloroso Muy doloroso Insoportable

9. Durante las últimas 4 semanas, ¿ha estado cojeando al andar debido a su cadera?

Rara vez/ nunca A veces, o solo al principio Con frecuencia, no solo al principio La mayoría del tiempo Todo el tiempo

10. Durante las últimas 4 semanas, ¿ha tenido algún dolor grave y repentino: “fulgurante”, “punzante” o “espasmos”, debido a la cadera afectada?

Ningún día Solo 1 o 2 días Algunos días La mayoría de días Todos los días

11. Durante las últimas 4 semanas, ¿en qué medida ha interferido su dolor debido a la cadera con su trabajo habitual (incluidas las tareas domésticas)?

Nada Un poco Moderadamente Mucho Totalmente

12. Durante las últimas 4 semanas ¿le ha molestado el dolor debido a la cadera en la cama por la noche?

Ninguna noche Solo 1 o 2 noches Algunas noches La mayoría de noches Todas las noches

© Oxford University Innovation Limited, 1998. All rights reserved.
Cualquier licencia para el uso de esta versión del Oxford Hip Score debe ser solicitada a Oxford University Innovation Ltd usando el siguiente link:
<https://innovation.ox.ac.uk/clinical-outcomes/patient-reported-outcome-measures/>

Anexo 2. Versión adaptada al español (España) del Oxford Knee Score.

1. Durante las últimas 4 semanas, ¿cómo describiría el dolor que tiene normalmente debido a su rodilla?

Ninguno	Muy leve	Leve	Moderado	Grave
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. Durante las últimas 4 semanas, ¿ha tenido algún problema para lavarse y secarse (todo el cuerpo) debido a su rodilla?

Ningún problema	Muy pocos problemas	Problemas moderados	Dificultad extrema	Imposible hacerlo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3. Durante las últimas 4 semanas, ¿ha tenido algún problema para subir y bajar de un coche o para usar el transporte público debido a su rodilla? (sea cual sea el que suele utilizar)

Ningún problema	Muy pocos problemas	Problemas moderados	Dificultad extrema	Imposible hacerlo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4. Durante las últimas 4 semanas, ¿durante cuánto tiempo ha podido andar antes de que el dolor debido a su rodilla se volviera grave (con o sin bastón)?

Sin dolor/Más de 30 minutos	16 a 30 minutos	5 a 15 minutos	Solo por casa	Nada/dolor grave al caminar
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5. Durante las últimas 4 semanas, tras una comida (sentado/a a una mesa), ¿en qué medida le ha resultado doloroso levantarse de la silla debido a su rodilla?

Nada doloroso	Ligeramente doloroso	Moderadamente doloroso	Muy doloroso	Insoportable
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6. Durante las últimas 4 semanas, ¿ha estado cojeando al andar debido a su rodilla?

Rara vez/ nunca	A veces, o solo al principio	Con frecuencia, no solo al principio	La mayoría del tiempo	Todo el tiempo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7. Durante las últimas 4 semanas, ¿podía arrodillarse y levantarse de nuevo después?

Caracterización de la Calidad de Vida Relacionada con la Salud en pacientes con artrosis de miembro inferior y validación de nuevas herramientas para su evaluación.

Sí, fácilmente	Con poca dificultad	Con dificultad moderada	Con extrema dificultad	No, imposible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8. Durante las últimas 4 semanas, ¿le ha molestado el dolor debido a la rodilla en la cama por la noche?

Ninguna noche	Solo 1 o 2 noches	Algunas noches	La mayoría de noches	Todas las noches
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9. Durante las últimas 4 semanas, ¿en qué medida ha interferido su dolor debido a la rodilla con su trabajo habitual (incluidas las tareas domésticas)?

Nada	Un poco	Moderadamente	Mucho	Totalmente
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10. Durante las últimas 4 semanas, ¿ha sentido que su rodilla podría no responderle o no soportar su peso repentinamente?

Rara vez/nunca	A veces, o solo al principio	Con frecuencia, no solo al principio	La mayoría del tiempo	Todo el tiempo
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

11. Durante las últimas 4 semanas, ¿podía hacer la compra para su casa usted solo/a?

Sí, fácilmente	Con poca dificultad	Con dificultad moderada	Con extrema dificultad	No, imposible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

12. Durante las últimas 4 semanas, ¿podía bajar un tramo de escaleras?

Sí, fácilmente	Con poca dificultad	Con dificultad moderada	Con extrema dificultad	No, imposible
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

© Oxford University Innovation Limited, 1998. All rights reserved.

Cualquier licencia para el uso de esta versión del Oxford Knee Score debe ser solicitada a Oxford University Innovation Ltd utilizando el siguiente link:

<https://innovation.ox.ac.uk/clinical-outcomes/patient-reported-outcome-measures/>

Anexo 3. Publicaciones

RESEARCH ARTICLE

Open Access

Cross-cultural adaptation and validation of the Spanish version of the Oxford Hip Score in patients with hip osteoarthritis



Jesús Martín-Fernández^{1,2,3*}, Pedro Gray-Laymón⁴, Antonio Molina-Siguero⁵, Javier Martínez-Martín⁶, Roberto García-Maroto⁷, Isidoro García-Sánchez⁸, Lidia García-Pérez^{9,10}, Vanesa Ramos-García⁹, Olga Castro-Casas¹ and Amaia Bilbao¹¹

Abstract

Background: Osteoarthritis (OA) of the hip is a disease that entails a major burden for patients and the society as a whole. One way of measuring this burden for the patient is through impact on Health-related Quality of Life (HRQL). The Oxford Hip Score (OHS) is a well-known tool to measure HRQL in patients with OA of the hip. This study aims to assess the psychometric properties of the Spanish-adapted version of the OHS, including its reliability, validity, and sensitivity to change.

Methods: Prospective observational study that included 361 patients diagnosed with hip OA (according to the criterion of the American College of Rheumatology) from 3 different Spanish regions. Their HRQL was assessed using a generic questionnaire, the EQ-5D-5 L, and two specific ones (the Western Ontario and McMaster Universities Osteoarthritis Index, WOMAC, and the OHS) adapted to Spanish. There was a follow-up period of 6 months, and the acceptability, psychometric properties, presence of ceiling and floor effects, validity, reliability, and sensitivity to changes of the OHS were measured.

Results: The OHS was fully answered in 99.4% of cases with no indication of ceiling or floor effects. Its factor structure can be explained in a single dimension. Its discriminative capacity was very good compared to the groups generated by the WOMAC and the EQ-5D-5 L. The correlation between the OHS and dimensions of the WOMAC or EQ-5D-5 L utilities was ≥ 0.7 . Excellent test-retest reliability (ICC = 0.992; CI95%: 0.994–0.998) and internal consistency (Cronbach's α = 0.928) were observed. The minimal clinically important difference (MCID) was 7.0 points, and the minimum detectable change (MDC) was 5.5 points. The effect size for moderate improvement in perceived HRQL was 0.73, similar to that of WOMAC dimensions and higher than the EQ-5D-5 L.

Conclusions: The Spanish-adapted version of the OHS is a useful, acceptable tool for the assessment of perceived HRQL in patients with hip OA, and has psychometric properties similar to those of the WOMAC that allow for discriminating both a patient's condition at a given moment and changes that can occur over time.

Keywords: Quality of life, Health status, Osteoarthritis, Hip, Validation studies

* Correspondence: jmferrandez@salud.madrid.org

¹ C^o Vilamanta (C.S. Navalcarnero), Gerencia Asistencial de Atención Primaria, Servicio Madrileño de Salud, Madrid, Spain

² Facultad de Ciencias de la Salud, Universidad Rey Juan Carlos, Madrid, Spain

Full list of author information is available at the end of the article



Background

Osteoarthritis (OA) is the most frequent joint disease, manifesting when structural changes in the joint cause pain and functional impairment. The prevalence of hip OA is high and is augmenting in developed countries due to increases in life expectancy and obesity pandemics [1, 2]. In a literature review by Pereira et al, the prevalence of hip OA was reported to be 10.9% (CI 95%: 10.6–11.2), although the figure was higher when based on radiological diagnosis rather than clinical evidence [3]. The prevalence of hip OA in Spain has been estimated to be 0.9% in the population >40 years of age [4], and 7.4% for people >60 years of age [5].

Hip OA greatly impacts the patient's perception of health-related quality of life (HRQL) [6], and entails a great burden for the individual and the society as a whole. Studies of international scope have estimated that OA of the knee and hip constitute 0.7% of all disability adjusted life years (DALY) [7]. The DALYs lost due to hip OA increased 60% between 1990 and 2010 [8]. In the USA, the yearly expenditure resulting directly from hip OA was calculated to be \$2827 per patient over 65 years of age (in the 1990s), and indirect costs can exceed that figure [1]. In Spain, the health-related expenses derived from OA can amount to 0.25–0.50% of the country GDP [9]. A study performed in Spain in 2007 estimated a yearly expenditure of €1500 per patient with hip or knee OA, 86% of which were direct costs [10].

It is necessary to incorporate the patient's self-perception of health condition to the study of chronic diseases such as OA, both for appraising their current condition and the results of interventions [11]. The HRQL is a measure of the patient's perception of their health condition that can be assessed via "generic" or "specific" tools. Generic tools are used to appraise health condition for any typology of patients, whereas specific tools are devised for a specific disease (e.g., OA of the hip), population segment (young vs. old), or type of problem (pain, dyspnea, *et cetera*) [12].

In the case of hip OA, there are several specific tools to evaluate HRQL, such as the Harris Score [13], the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) for the assessment of OA of lower limbs [14], the Hip disability and Osteoarthritis Outcome Score (HOOS) for patients undergoing conservative treatment or surgery, the Hip Outcome Score [15] for patients about to undergo arthroscopy, or the Oxford Hip Score (OHS). Of them, only the WOMAC [16] and the Hip Outcome Score are validated in Spanish [17].

Although the OHS has been adapted to Spanish, its psychometric properties have not been assessed in the Spanish population setting. The OHS was designed to appraise the impact of total hip replacement surgery and was found to be more accurate than other generic questionnaires for that purpose [18]. Owing to its good psychometric properties, it has been favorably compared

to other widely used tools that are more difficult to administer [19]. It has been adapted to Dutch [20], French [21], German [22], Italian [23], Danish [24], Turkish [25], and several Asian languages [26–28]. The questionnaire was also adapted to Spanish in Colombia and partially validated, although neither its sensitivity to changes nor factorial structure were checked [29]. Of all the mentioned adaptations, the factorial structure has only been validated by means of a confirmatory factor analysis in the original version of the OHS.

This study aims to assess the psychometric properties of the Spanish-adapted version of the OHS, including its factorial structure and other aspects of reliability, validity, and capacity to detect changes.

Methods

Design

Prospective, observational study, with a follow-up period for the recruited subjects of 6 months.

Sampling and sample size

Opportunistic sampling was performed. Patients diagnosed with hip OA according to the criterion of the American College of Rheumatology [30] were recruited from traumatology, rheumatology, and primary care consultations in Vizcaya, Madrid, and Tenerife. Subjects were included in a consecutive way between January and December 2015. Not understanding Spanish, illiteracy, or being diagnosed with cognitive impairment were considered to be exclusion criteria.

Sample size was calculated on the basis of the most stringent analysis method employed: three hundred patients were estimated necessary to perform a confirmatory factorial analysis (CFA) using a single questionnaire that comprised 12 items [31]. This sample size allowed for estimating intraclass correlation coefficients (ICC) >0.8 with precision values $\geq 10\%$ [32].

Variables

The following personal characteristics were recorded for all participants: age, gender, body mass index (BMI), arthritis-affected joints, previous joint replacement surgery, and comorbidity, measured using the Charlson's index [33]. In order to evaluate self-perception of HRQL, patients completed 3 questionnaires in their Spanish-adapted version: a generic one, the EQ-5D with a 5-level scale (EQ-5D-5 L) [34], and two specific ones, the WOMAC [14] and the OHS [18].

The EQ-5D-5 L [34] inquires about current self-perceived health condition and comprises two parts. The first part includes 5 questions on mobility, self-care, performance of daily-life activities, pain/discomfort, and anxiety/depression; each dimension is measured on a scale from 1 to 5; a single weighted score, called the

utility index, is then obtained from these 5 questions, so that the greater the score the better the health condition [35]. The second part consists of a visual analogue scale (VAS) ranging from 0 (worst health condition) to 100 (best imaginable health condition).

The WOMAC [14] is a self-completed questionnaire, specifically aimed at patients suffering from OA of the hip or knee. Its multidimensional scale comprises 24 items clustered according to 3 dimensions: pain (5 items), stiffness (2 items), and physical functionality (17 items). A Likert-type scale was used with 5 possible answers to account for the intensity of each item (none, slight, moderate, severe, extreme), so each item receives a score from 0 to 4. The scores are then summed and standardized from 0 to 100 (best to worst), so that the greater the score the worse the patient's health condition. This questionnaire has been adapted and validated for the Spanish setting [16].

The OHS is a self-administered questionnaire that can be completed via a personal interview or mailed by the patient after completion. It comprises 12 questions, with 5 possible answers each, to assess the patient's perception of quality of life during the last 4 weeks. It has been employed with patients suffering from hip OA, both to study their baseline condition and to evaluate changes following prosthetic implant. Each answer receives a score from 0 to 4, where 4 is the best possible outcome [36]. A total score is obtained from the sum of all answers, ranging from 0 to 48, where 48 is the best possible outcome. The Spanish-adapted version was developed by performing a translation and linguistic validation using protocols consistent with internationally recognized good-practice guidelines under agreement with Oxford University Innovation™ (see Additional file 1).

The recruited participants from Madrid were interviewed 7 to 14 days after the inclusion visit and, after verifying that there were no changes in their health condition, the OHS was repeated to check test-retest reliability.

All included patients were interviewed after a 6 months follow-up period: they were asked if they had undergone replacement surgery, all the questionnaires were repeated (EQ-5D-5 L, WOMAC, and OHS), and transition questions were posed to check for changes in their perception of global health.

Statistical analysis

Continuous variables are described by their central tendency and dispersion, whereas qualitative variables are expressed by their percentages. Confidence intervals were set at 95% (CI 95%).

Acceptability and ceiling and floor effects

The number of non-completed questionnaires and unanswered questions was recorded.

Ceiling or floor effects were considered to be present if more than 15% of respondents reported the highest or lowest possible score, respectively [37].

Analysis of psychometric properties

Validity

The construct validity was assessed via an exploratory factor analysis (EFA). Bartlett's test of sphericity and a Kaiser-Meyer-Olkin (KMO) test were performed to evaluate the adequacy of employing such analysis. The null hypothesis of Bartlett's test states that the matrix of observed correlations is a singular matrix. Rejecting the null hypothesis allows for confirming the existence of linear relationships between factors and the explanatory variable. The KMO sampling adequacy test provides a measure of the variance among variables, and values >0.90 are considered optimal [38]. Factor loadings were calculated, with values >0.40 considered to be optimal, and so were communalities, that express the percentage of the item's variance explained by each of the studied factors.

In order to complement our results and confirm the hypothesis of unidimensionality of the questionnaire, a confirmatory factor analysis (CFA) for categorical variables was carried out. The robust unweighted least squares estimator was used and the following fit indices were calculated [39, 40]: the root mean square error of approximation (RMSEA), for which a value <0.08 was acceptable, and the Tucker-Lewis Index (TLI) and Comparative Fit Index (CFI), both of which had to be >0.95 to be considered satisfactory [41]. Additionally, factor loadings were examined and those ≥ 0.40 were considered acceptable. The model was considered satisfactory if it surpassed these acceptability criteria.

The validity of the known groups was appraised by comparing the scores obtained in the OHS with each tercile of the EQ-5D-5 L and WOMAC distributions.

Convergent validity was assessed by calculating Pearson's r or Spearman's ρ , which were then used to find the correlations between the scales of the OHS and those of the WOMAC and the EQ-5D-5 L. A threshold of 0.7 was set [37] for associations to be considered strong.

Reliability

Internal consistency was assessed by calculating Cronbach's α [42] for the scores obtained at the inclusion visit. This coefficient accounts for internal correlations of all items in a scale, so the greater Cronbach's α is (range 0.0 to 1.0), the greater the consistency of the scale and the greater the probability that the questionnaire underlies a single dimension. In the case of a unidimensional tool comprising 12

items, Cronbach's α is required to be >0.85 for the internal consistency to be considered optimal [43].

The test-retest reliability was checked using the ICC for comparing the scores of the test with the retest in the sub-sample from Madrid. According to the suggested classification for different reliability measurements [44], ICC values >0.7 are considered to be acceptable and >0.9 optimal.

Sensitivity to change

The OHS questionnaire was repeated 6 months after the inclusion visit in order to evaluate its capacity to detect changes in the evolution of the disease. Transition questions were posed that inquired about the change in the hip condition perceived by the patient relative to the 6 previous months. Five possible answers (much worse, slightly worse, equal, slightly better, much better) were given and recorded on a scale. These questions aimed at appraising the sensitivity to change of the OHS questionnaire. Transition questions for the WOMAC were answered on the same scale, but they were specific for each of its domains (pain, stiffness, and physical functionality). Correlations between score changes in HRQL questionnaires and transition questions were assessed by calculating Spearman's rho.

Changes in the OHS and EQ-5D-5 L questionnaires were calculated by subtracting initial from final scores, so positive values indicated an improvement in general condition. This was the opposite for the WOMAC, where final scores were subtracted from initial ones, and therefore positive values also indicated improvement. For each group of patients, baseline scores were then compared to those obtained at the follow-up period of 6 months to check if significant changes had occurred according to the transition questions. For every observed change, the effect size (ES) was calculated as the ratio between the mean and standard deviation (SD). Changes were considered to be moderate for values >0.5 , and large for >0.8 [45]. Obtained values of ES were then compared to those of the WOMAC and EQ-5D-5 L scales. Responsiveness parameters were also estimated separately for patients who had suffered a hip arthroplasty and those who did not.

Additionally, the minimal clinically important difference (MCID) and the minimal detectable change (MDC) were estimated. These two measures correlate with responsiveness, but are more clinically oriented and focused at the individual level. The MCID was calculated using the mean change of patients that reported moderate improvement (feeling "slightly better") at 6 months of the follow-up [46].

The MDC expresses the minimal magnitude of change above which the observed change is likely to be real and

not just measurement error. The standard error of measurement (SEM), which represents the amount of error associated with a particular individual's assessment, was estimated as the square root of the mean square error term from the ANOVA [47, 48]. From the SEM, the MDC was derived as follows [37, 47].

$$\text{MDC} = \text{SEM} \times z\text{-score} \times \sqrt{2}$$

A 95% confidence level ($\text{MDC}_{95\%}$) was established, corresponding to a z-value of 1.96. The interpretation of $\text{MDC}_{95\%}$ is that if a patient has a change score equal to or higher than the $\text{MDC}_{95\%}$ threshold, it is possible to state with 95% confidence that this change is reliable and not the result of measurement error. Finally, to determine if the MCID surpassed the $\text{MDC}_{95\%}$, MCID was divided by the $\text{MDC}_{95\%}$ [49] so that if this ratio exceeded 1, the MCID could be discriminated from measurement error.

All effects were considered statistically significant at $p < 0.05$.

Results

The study included 361 subjects: 157 from Vizcaya, 124 from Madrid, and 80 from Tenerife. Patients were recruited at primary care (37.7%), traumatology (46.5%), and rheumatology (10.8%) consultations. Women comprised 53.2% (CI 95%: 48.0–58.4%) of the sample and the average age was 67.8 years (CI 95%: 66.7–69.1 years).

Replacement of the contralateral hip had occurred in 17.5% (CI 95%: 13.6–21.4%) of cases. Charlson's index had an average value of 0.8 points (CI 95%: 0.7–1.0), and mean BMI was 28.2 (CI 95%: 27.7–28.6).

Table 1 shows the outcome expressed by patients for the OHS, WOMAC, and EQ-5D-5 L questionnaires.

Acceptability and ceiling and floor effects

The obtained data allowed for summarizing the outcome of the OHS questionnaire in 359 cases (99.4%; CI95%: 98.7–100%). Questions 3 and 4 were answered in all cases, and questions 2, 5, 7, 8, and 10 in all but one. Questions 2, 6, and 11 were not answered in 2 occasions, and questions 9 and 12 in 3 occasions. The possible responses were the same for all questions, and so was the possible score range (0 to 4). There was no question for which more than 35% of the responses were concentrated at the top or lowest end of the scale: question 6 obtained the greatest percentage of responses for the lowest score (27.7%), and question 9 for the highest possible score (33.8%). For the total score, there was no aggregation at the low end of the scale and only 0.84% of the responses scored 48 out of 48 possible points in the inclusion visit. This score was reached by 3.08% of the patients that underwent hip replacement in the visit after six months. Hence, the presence of floor or ceiling effects was ruled out.

Table 1 Scores from the OHS, WOMAC, and EQ-5D-5 L questionnaires

	n	Mean score (CI 95%)	Median (interquartile range)
OHS	359	22.8 (21.7–23.9)	22.0 (15.0–30.0)
WOMAC pain	360	45.8 (43.4–48.2)	45.0 (30.0–60.0)
WOMAC stiffness	360	48.3 (45.6–51.0)	50.0 (25.0–62.5)
WOMAC physical functionality	360	52.4 (50.0–54.8)	51.5 (36.8–70.6)
EQ-5D-5 L utility	357	0.52 (0.49–0.55)	0.61 (0.32–0.73)
EQ-5D-5 L VAS	357	54.5 (52.2–56.8)	50.0 (40.0–70.0)

Utility index score 0 to 1, where 0 = condition comparable to death, and 1 = perfect health condition. Negative scores are allowed

OHS Oxford Hip Score. Score range 1 to 48, the higher the score the better, WOMAC Western Ontario and McMaster Universities Osteoarthritis Index. Scale range 1 to 100, the higher the score the worse the condition

VAS Visual Analogue Scale. Score range 1 to 100, where 0 = worst possible health condition, and 100 = best imaginable health condition

Validity

In order to study the validity of the construct, an EFA was performed and a unidimensional structure was revealed with a single factor that explained 55.5% of variance (KMO = 0.945, Bartlett's test of sphericity $\chi^2 = 2667$, 66 degrees of freedom, $p < 0.001$). Every factor loading was > 0.50 , and communalities were ≥ 0.40 with the exception of questions 6 and 10 (Table 2).

Fit indices resulting from the performed CFA were adequate (Fig. 1): (a) the RMSEA was 0.082, a very close value to the set threshold of 0.08; and (b) the CFI and TLI were 0.982 and 0.977, respectively, both exceeding the benchmark of 0.95. All factor loadings were statistically significant ($p < 0.001$), ranging from 0.57 to 0.88.

Validity of known groups, which is a measure of the questionnaire's discriminatory capacity, can be observed in Table 3. It shows average scores of the OHS with their CI95% for each tercile of the distributions of the WOMAC and EQ-5D-5 L scales. Differences between the 3 groups are clearly shown by the OHS scores, with average changes of 4.7 to 12.0 points per tercile.

Table 4 shows the correlations between the OHS scores and the WOMAC domains or the EQ-5D-5 L domains, utility index and VAS. Given the different types of scale measures, negative correlations with the WOMAC and positive ones with the EQ-5D-5 L were to be expected. All associations were strong except for the stiffness scale of the WOMAC questionnaire, where the correlation was at the limit of the set threshold, and the EQ-5D-5 L VAS.

Correlations between the scores of the WOMAC scales on pain, stiffness, and physical functionality and the EQ-5D utilities were -0.769 , -0.628 , and -0.829 , respectively (all values were statistically significant, $p < 0.001$). Correlations between the scores of the WOMAC scales on pain, stiffness, and physical functionality and the EQ-5D VAS were -0.563 , -0.410 , and -0.560 , respectively (all values were statistically significant, $p < 0.001$).

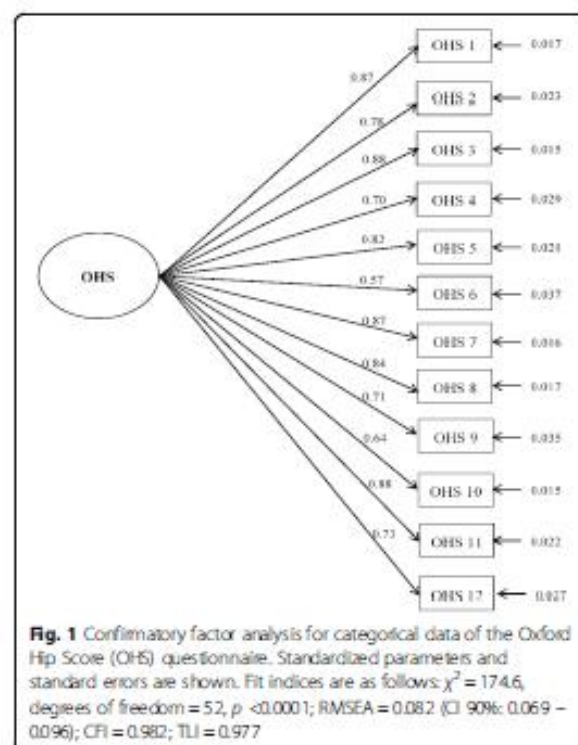
Reliability

Internal consistency was assessed via Cronbach's α , which was calculated to be 0.928 for the OHS questionnaire. For the 124 subjects that repeated the questionnaire 7 to

Table 2 Exploratory factor analysis of the Oxford Hip Score (OHS) items

Question: During the past 4 weeks...	Exploratory factor analysis 1 factor	
	Factor loading	Communality
During the past 4 weeks... How would you describe the pain you usually have from your hip? (OHS 1)	0.814	0.662
Have you had any trouble with washing and drying yourself (all over) because of your hip? (OHS 2)	0.791	0.626
Have you had any trouble getting in and out of a car or using public transport because of your hip? (OHS 3)	0.842	0.709
Have you been able to put on a pair of socks, stockings or tights? (OHS 4)	0.738	0.545
Could you do the household shopping on your own? (OHS 5)	0.797	0.635
For how long have you been able to walk before pain from your hip becomes severe? (OHS 6)	0.558	0.311
Have you been able to climb a flight of stairs? (OHS 7)	0.834	0.696
After a meal (sat at a table), how painful has it been for you to stand up from a chair because of your hip? (OHS 8)	0.810	0.656
Have you been limping when walking, because of your hip? (OHS 9)	0.698	0.488
Have you had any sudden, severe pain - 'shooting', 'stabbing' or 'spasms' - from the affected hip? (OHS 10)	0.630	0.396
How much has pain from your hip interfered with your usual work (including housework)? (OHS 11)	0.852	0.726
Have you been troubled by pain from your hip in bed at night? (OHS 12)	0.718	0.515

OHS Oxford Hip Score



14 days after their inclusion in the study, ICC was 0.992 (CI 95%: 0.994–0.998).

Sensitivity to change

A follow-up on 313 subjects took place after 6 months. Of them, 65 had undergone hip replacement surgery and 94 (30.0%, CI 95%: 25.0–35.1%) reported feeling “slightly better” or “much better” on the side of the hip for which they entered the study. Of the follow-up sample, 133 (42.5%; CI 95%: 37.0–48.0%) stated feeling “slightly worse” or “much worse”.

Table 5 shows the mean change in the scores obtained from the employed questionnaires when the patient had perceived a change in their health condition. First, the correlations between score changes between the HRQL

questionnaires and transition questions were assessed. The correlation between the change in the OHS score and transition questions was moderate (Spearman’s rho = 0.636, $p < 0.0001$). The correlations between changes in the domains of the WOMAC and specific transition questions were also moderate (Spearman’s rho absolute value between 0.544 and 0.635; $p < 0.0001$).

The ES of the change in the OHS was 0.73 for subjects that reported feeling “slightly better” and 1.71 for those that reported feeling “much better”. Sensitivity to change obtained lower ES values for negative changes, with values of 0.42 and 0.69 in the case of subjects reporting “slightly worse” and “much worse”, respectively. A clear gradient in the scores was observed that depended on the change perceived by the patient, which was significantly different for those feeling “slightly worse”, “slightly better”, and “much better”. There was a small overlap between subjects feeling “much worse” and “slightly worse”. The OHS questionnaire proved to be a more sensitive tool than the EQ-5D-5 L, and similar to the WOMAC.

Table 6 shows the mean change in the scores obtained from the OHS questionnaires for both patients that had undergone hip arthroplasty and those who did not. Results were consistent with those from the whole sample, although improvements perceived by patients who underwent hip arthroplasty were significantly greater.

The average change in the OHS scores was 7.0 points (SD = 9.6) in the case of subjects that felt moderate improvement, which was the figure used for calculating the MCID. The SEM was calculated to be 2.0 and hence, the estimated value of $MDC_{95\%}$ was 5.5. The obtained ratio $MCID/MDC_{95\%}$ was 1.3.

Discussion

The Spanish version of the OHS is a valid tool for measuring HRQL in patients suffering from hip OA, and is both reliable and sensitive to changes. Additionally, it is very well accepted by the population it addresses, as proven by the extraordinarily high response rate, although in this case it could be influenced by the way in which it was administered, in the clinical setting.

Table 3 Average scores of the OHS. Patients are classified according to the tertiles obtained from the WOMAC and EQ-5D-5 L questionnaires

	OHS score for the lower tertile of the distribution (CI 95%)	OHS score for the middle tertile of the distribution (CI 95%)	OHS score for the top tertile of the distribution (CI 95%)
WOMAC pain	32.2 (30.7–33.6)	22.7 (21.1–24.2)	13.7 (12.6–14.9)
WOMAC stiffness	29.7 (28.1–31.2)	22.1 (20.8–23.4)	14.4 (13.0–15.7)
WOMAC physical functionality	33.6 (32.3–34.9)	21.6 (20.6–22.6)	13.0 (11.9–14.1)
EQ-5D-5 L utility	12.7 (11.7–13.7)	22.4 (21.2–23.5)	32.4 (31.1–33.9)
EQ-5D-5 L VAS	15.5 (14.0–17.0)	24.0 (22.3–25.7)	28.7 (26.7–30.5)

OHS Oxford Hip Score, WOMAC Western Ontario and McMaster Universities Osteoarthritis Index, VAS Visual Analogue Scale

Table 4 Correlations between the OHS score and the WOMAC scales or the EQ-5D-5 L

	OHS	
	r/rho	p-value
WOMAC		
Pain	-0.807 ^a	<0.001
Stiffness	-0.686 ^a	<0.001
Functionality	-0.893 ^a	<0.001
EQ-5D-5 L		
Mobility,	-0.768 ^b	<0.001
Self-care,	-0.728 ^b	<0.001
Performance of daily-life activities	-0.748 ^b	<0.001
Pain/discomfort	-0.778 ^b	<0.001
Anxiety/depression	-0.562 ^b	<0.001
Utility Index	0.835 ^a	<0.001
VAS	0.575 ^a	<0.001

Higher scores in the OHS and EQ-5D-5 L indicate better health condition, and the opposite happens in the case of the WOMAC

OHS Oxford Hip Score, WOMAC Western Ontario and McMaster Universities Osteoarthritis Index

^a Pearson's r

^b Spearman's rho

The validity of the OHS has been assessed from different perspectives, although apparent validity was not one of them given it is an adaptation.

The validity of known groups, namely its discriminatory validity, appears to be adequate since scores differ between subjects classified according to their HRQL, via specific or general questionnaires. Significant ceiling or floor effects that could compromise such discriminatory capacity were not found. The presence of floor effect has not ever reported with the OHS. On the contrary, certain studies have observed a ceiling effect in postoperative patients [24], although the majority have not [21, 22, 50]. The previously shown results rule out the presence of ceiling effect, even in patients who had undergone hip replacement, in agreement with the results of the original version [51].

The analysis of convergent validity showed correlations with the specific scales of the WOMAC and the generic scales of the EQ-5D-5 L. Such correlations were stronger than those found between the original questionnaire and generic tools for measuring HRQL [51], and similar or slightly stronger than adapted versions of the OHS to other languages, such as German [22] or Dutch [20].

The construct validity was also part of the validation. The factorial structure of the OHS has been previously discussed, and several authors have proposed to differentiate 2 domains within it: pain and physical functionality [52]. When attempting to check if a single- or double-factor structure worked better, the outcome supported both possibilities, although there were several items that

saturated both factors when considering a bidimensional structure [53]. For these reasons and in view of the outcome of the performed EFA, which was similar to other adapted versions [54], a unidimensional structure was tested, which seemed a correct approach for this adaptation given that the values obtained in the CFA were close to the acceptability threshold for the RMSEA and optimal for the TLI and CFI [41].

Cronbach's α , which accounts for internal consistency, was better than for the original scale at the inclusion visit (0.93 vs 0.84) [55]. Although a very high value of Cronbach's α could indicate that the items are redundant, this is unlikely to be the case, since it was ruled out by the factorial analysis. This coefficient is useful for estimating reliability, particularly for a unidimensional test. If a test shows a high value of α , then it can be concluded that its variance is largely attributable to general and group factors. When the existence of a single factor has been demonstrated, then Cronbach's α can be used to conclude that the set of items is unidimensional [43].

Test-retest reliability was measured via ICCs and found to be excellent, with values >0.90 that in a sample of 124 patients allows for classifying the tool as reliable [37]. Reported values of ICC were slightly higher than those found in other studies (range from 0.89 to 0.97) [22–24, 27, 50], which may be due to the way in which the OHS score was obtained in the follow-up, namely by telephone interview.

The reliability study of this adapted version of the OHS yielded values of internal consistency and reliability that were similar to another Spanish-validated tool, the Hip Outcomes Score, which is also designed to appraise changes in perceived HRQL by patients following hip surgery [17].

The discriminatory capacity of the questionnaire, which accounts for its potential to discriminate patients in different situations, was satisfactory; however, the tool has also proven its usefulness to study the subject's perception of change in their own situation, that is, its evaluative capacity is adequate [11]. The instrument was originally designed for this purpose and this study has confirmed the potential of its adaptation to Spanish, not only in patients that undergo hip surgery but also in the short-term evolution of a cohort of patients suffering from hip OA.

The ES for "moderate" positive changes showed values that were slightly under the set threshold of 0.8 points, and similar to those of the WOMAC. The ES for changes following surgery was 1.93 in the validation process of the original questionnaire [56], which is only comparable with improvements in patients that underwent hip replacement (ES = 1.35). The OHS proved to be superior to other questionnaires, like the WOMAC and the generic EQ-5D-5 L, when assessing significant changes, as is the case of hip replacement [51]. In this work, the OHS showed a similar

Table 5 Changes observed in the OHS, EQ-5D-5 L, and WOMAC questionnaires at 6 months of the follow-up period, in patients that reported health condition changes

	The hip condition is "much worse" N = 53		The hip condition is "slightly worse" N = 78		The hip condition is "slightly better" N = 34		The hip condition is "much better" N = 60	
	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.
OHS	-5.8 (-8.1-3.4)	0.69	-2.4 (-3.6-1.1)	0.42	7.0 (3.7-10.4)	0.73	18.7 (15.9-21.6)	1.71
EQ-5D-5 L utility	-0.16 (-0.23-0.08)	0.57	-0.06 (-0.11-0.01)	0.09	0.19 (0.10-0.28)	0.73	0.40 (0.33-0.47)	1.48
EQ-5D-5 L VAS	-5.3 (-12.9-2.2)	0.20	-1.6 (-5.6-2.4)	0.09	9.0 (0.6-17.3)	0.37	20.6 (14.7-26.5)	0.91
	Pain is "much worse" N = 46		Pain is "slightly worse" N = 77		Pain is "slightly better" N = 42		Pain is "much better" N = 60	
WOMAC pain	Average change (CI 95%) -13.5 (-18.5-8.4)	E.S. 0.79	Average change (CI 95%) -4.1 (-7.2-1.0)	E.S. 0.30	Average change (CI 95%) 1.49 (8.5-21.3)	E.S. 0.73	Average change (CI 95%) 38.3 (32.2-44.3)	E.S. 1.64
	Stiffness is "much worse" N = 38		Stiffness is "slightly worse" N = 82		Stiffness is "slightly better" N = 33		Stiffness is "much better" N = 56	
WOMAC Stiffness	Average change (CI 95%) -12.2 (-20.9-3.4)	E.S. 0.50	Average change (CI 95%) -6.9 (-11.5-6.2)	E.S. 0.32	Average change (CI 95%) 1.78 (8.5-27.1)	E.S. 0.68	Average change (CI 95%) 38.6 (31.1-46.1)	E.S. 1.37
	Physical functionality is "much worse" N = 48		Physical functionality is "slightly worse" N = 82		Physical functionality is "slightly better" N = 33		Physical functionality is "much better" N = 47	
WOMAC Physical functionality	Average change (CI 95%) -8.7 (-14.5-2.9)	E.S. 0.43	Average change (CI 95%) -4.9 (-8.4-1.5)	E.S. 0.31	Average change (CI 95%) 21.3 (14.0-28.7)	E.S. 1.03	Average change (CI 95%) 43.2 (36.2-50.2)	E.S. 1.81

Final scores were subtracted from basal scores in the OHS and EQ-5D-5 L, and the opposite happened in the WOMAC. This way, positive outcomes always indicate improvement

E.S. Effect Size, OHS Oxford Hip Score, WOMAC Western Ontario and MacMaster Universities Osteoarthritis Index

Table 6 Changes observed in the OHS questionnaires at a 6 months follow-up for patients having undergone hip arthroplasty or not

Not Arthroplasty	The hip condition is "much worse" N=51		The hip condition is "slightly worse" N=76		The hip condition is "slightly better" N=19		The hip condition is "much better" N=14	
	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.
OHS	-5.8 (-8.1--3.4)	0.69	-2.4 (-3.6--1.1)	0.42	3.6 (-0.9-8.0)	0.39	12.4 (6.8-17.9)	1.28
Arthroplasty	The hip condition is "much worse" N=1		The hip condition is "slightly worse" N=2		The hip condition is "slightly better" N=15		The hip condition is "much better" N=46	
OHS	-2.00	-	-3.5 (-6.7--53.7)	0.55	11.4 (6.7-11.1)	1.35	20.7 (17.5-23.8)	1.94

Final scores were subtracted from basal scores
E.S. Effect Size, OHS Oxford Hip Score

capacity for detecting "moderate" changes compared to the WOMAC, but was slightly better when examining "significant" changes.

For subjects that reported feeling a "moderate" improvement in their condition, MCID was 7.0 points. For the original version of the OHS, MCID was calculated from the scores reported by patients that had undergone hip replacement surgery, and values of 7.5 points were obtained [57]. Using the criterion of estimating MCID as half the SD of the distribution of scores given by subjects that had experienced changes [58], MCID can be calculated to be around 4.8 points. According to this criterion, the values of MCID for the original version would be between 3 and 5 points, similarly to those obtained in our study [36].

In agreement with other studies, the evaluative capacity was greater for detecting positive rather than negative changes [59], although the capacity observed for the OHS to detect negative changes was similar or greater than the WOMAC and, of course, than that of generic questionnaires like the EQ-5D-5 L.

The MDC_{95%} was 5.5 points as calculated from the SEM, a value that is similar to the original questionnaire (MDC_{90%} = 4.85 points) [57]. The MDC represents the lowest score change (at the particular patient level) that is not the result of measurement error of the instrument. The MDC is based on the standard error of measurement, which depends on the accuracy and variability of its components [47], and can be understood as the lowest bound of real change, although it may not indicate clinical significance. The ratio between the MCID and MDC_{95%} was higher than 1, indicating that the MCID can be discriminated clearly from measurement error.

There are some limitations to this work. The studied sample may not be representative of the Spanish population, despite including patients from different geographic regions and at various stages of the disease evolution. On the other hand, the used methodology (the classical test theory), with its assumptions and constraints, entails certain limitations to evaluate psychometric properties; in

order to overcome them, the validation process has been complemented by performing a CFA specific for categorical data, which employs statistical analysis to validate *a priori* made assumptions [39].

Traditionally, the OHS has been used to assess the impact of hip replacement surgery [60] on HRQL, as well as other surgical [50] or non-surgical [23] procedures. Few studies have focused on its discriminatory capacity [25, 29]. This work highlights the discriminatory capacity of the tool and appraises its sensitivity to changes in the general evolution of the disease, including both patients that undergo joint replacement and those who do not. Its usefulness is similar to other instruments having a broad experience of use after being adapted to other languages, like the WOMAC, and displays greater capability to detect changes than generic tools like the EQ-5D-5 L.

Conclusions

The Spanish adaptation of the OHS is a useful instrument to assess perception of HRQL in patients suffering from hip OA, being well-accepted, and with good psychometric properties that support its use for evaluating a patient's condition at a given moment, and for appraising changes over time.

Incorporating this kind of tools to usual clinical practice will facilitate the valid and reliable evaluation of a patient's self-perceived health condition and the outcome of interventions, both at the individual and population level.

Additional file

Additional file 1: Adapted version of the Oxford Hip Score - Spanish (Spain). (DOC 47 kb)

Abbreviations

BMI: Body mass index; CFA: Confirmatory factor analysis; CFI: Comparative fit index; CI95%: 95% confidence interval; EFA: Exploratory factor analysis; ES: Effect size; GDP: Gross domestic product; HRQL: Health-related quality of life; ICC: Intraclass correlation coefficient; KMO: Kaiser-Meyer-Olkin; MCID: Minimal clinically important difference; MDC: Minimal detectable change; OA: Osteoarthritis; OHS: Oxford hip score; RMSEA: Root mean square error of approximation; SEM: Standard error of measurement; TL: Tucker-Lewis index;

VAS: Visual analogue scale; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index

Acknowledgements

This study has been financed by the Instituto de Salud Carlos III and the FEDER (European Regional Development Fund) (PI1300560, PI1300518, and PI1300648). We would like to thank Oxford University Innovation for providing us with the adapted version of the Oxford Hip Score - Spanish (Spain).

Funding

This study has been financed by the Instituto de Salud Carlos III and the FEDER (European Regional Development Fund) (PI1300560, PI1300518 y PI1300648).

Availability of data and materials

The datasets used or analyzed during the current study are available from the corresponding author upon reasonable request.

Authors' contributions

JMF, RGM, LGP, VRG and AB conceived and designed the experiments. JMF, PGL, AMS, JMM, IGS and OCC performed the experiments. JMF and AB analyzed the data. JMF, PGL, AMS, JMM, RGM, IGS, LGP, VRG, OCC and AB discussed the results. JMF and AB drafted the manuscript. JMF, PGL, AMS, JMM, RGM, IGS, LGP, VRG, OCC and AB revised and approved the manuscript.

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

All included patients provided written consent to participate in the study and approval was granted by all the relevant Ethics Committees for Clinical Research (CEIC de Euzkadi, Hospital Fundación Jiménez Díaz, Hospital Universitario de Fuenlabrada, Hospital Universitario Fundación Alcorcón, Hospital Universitario de Canarias, Hospital Universitario Nuestra Señora de Candelaria).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹CI Villamarta (C.S. Navacerrada). Gerencia Asistencial de Atención Primaria. Servicio Madrileño de Salud, Madrid, Spain. ²Facultad de Ciencias de la Salud, Universidad Rey Juan Carlos, Madrid, Spain. ³Red de Investigación en Servicios Sanitarios y Enfermedades Crónicas (REDISSEC), Madrid, Spain. ⁴Servicio de Cirugía Ortopédica y Traumatología, Hospital Universitario Rey Juan Carlos, Madrid, Spain. ⁵C.S. Presentación Sabido, Gerencia Asistencial de Atención Primaria, Servicio Madrileño de Salud, Madrid, Spain. ⁶Servicio de Cirugía Ortopédica y Traumatología, Hospital Universitario Fundación Alcorcón, Madrid, Madrid, Spain. ⁷Servicio de Cirugía Ortopédica y Traumatología, Hospital Universitario Clínico San Carlos, Servicio Madrileño de Salud, Madrid, Spain. ⁸Servicio de Traumatología y Cirugía Ortopédica, Hospital Galdakao-Usansolo (Osakidetza), Galdakao (Bizkaia), Spain. ⁹Fundación Canaria de Investigación Sanitaria (FUNICANIS), Santa Cruz de Tenerife, Spain. ¹⁰Red de Investigación en Servicios Sanitarios y Enfermedades Crónicas (REDISSEC), Santa Cruz de Tenerife, Spain. ¹¹Unidad de Investigación, Hospital Universitario Basurto (Osakidetza) - Red de Investigación en Servicios de Salud en Enfermedades Crónicas (REDISSEC), Bilbao (Bizkaia), Spain.

Received: 18 March 2017 Accepted: 10 May 2017

Published online: 22 May 2017

References

- Nho SJ, Kymes SM, Callaghan JJ, Felson DT. The burden of hip osteoarthritis in the United States: epidemiologic and economic considerations. *J Am Acad Orthop Surg*. 2013;21 Suppl 15:1–6.
- Ackermann IN, Osborne RH. Obesity and increased burden of hip and knee joint disease in Australia: results from a national survey. *BMC Musculoskelet Disord*. 2012;13:254.
- Pereira D, Peleteiro B, Araujo J, Branco J, Santos RA, Ramos E. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. *Osteoarthr Cartil*. 2011;19:1270–85. Elsevier Ltd.
- Prieto-Alhambra D, Judge A, Javald MK, Cooper C, Diez-Perez A, Arden NK. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: influences of age, gender and osteoarthritis affecting other joints. *Ann Rheum Dis*. 2014;73:1659–64.
- Quintana JM, Arostegui I, Escobar A, Azkarate J, Goenaga JI, Lafuente I. Prevalence of knee and hip osteoarthritis and the appropriateness of joint replacement in an older population. *Arch Intern Med*. 2008;168:1576–84.
- van der Waal JM, Tenwee CB, van der Windt DAWM, Bouter LM, Dekker J. The impact of non-traumatic hip and knee disorders on health-related quality of life as measured with the SF-36 or SF-12. A systematic review. *Qual Life Res*. 2005;14:1141–55.
- Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis*. 2014;73:1323–30.
- Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380:2163–96.
- Puig-Junoy J, Ruiz Zamora A. Socio-economic costs of osteoarthritis: a systematic review of cost-of-illness studies. *Semin Arthritis Rheum*. 2015;44:531–41.
- Loza E, Lopez-Gomez JM, Abesoia L, Maese JJ, Carmona L, Balle-Gualda E. Economic burden of knee and hip osteoarthritis in Spain. *Arthritis Rheum*. 2009;61:158–65.
- Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med*. 1993;118:622–9.
- McHorney CA. Health status assessment methods for adults: past accomplishments and future challenges 1. *Annu Rev Public Health*. 1999;20:309–35. Annual Reviews 4139 B Camino Way, PO Box 10139, Palo Alto, CA 94303–0139.
- Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. *J Bone Jt Surg Am*. 1969;51:737–55. The American Orthopedic Association.
- Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol*. 1988;15:1833–40.
- Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the hip outcome score. *Arthroscopy*. 2006;22:1304–11.
- Escobar A, Quintana JM, Bilbao A, Azkarate J, Goenaga JI. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. *Western Ontario and McMaster Universities Osteoarthritis Index*. *Clin Rheumatol*. 2002;21:466–71.
- Seijas R, Salient A, Ruiz-Ibán MA, Ares O, Marín-Peña O, Cuéllar R, et al. Validation of the Spanish version of the hip outcome score: a multicenter study. *Health Qual Life Outcomes*. 2014;12:70.
- Dawson J, Fitzpatrick R, Murray D, Carr A. Comparison of measures to assess outcomes in total hip replacement surgery. *Qual Saf Heal Care*. 1996;5:81–8.
- Kálarajah Y, Aruzza K, Hulme C, Molloy S, Drabu KJ. Health outcome measures in the evaluation of total hip arthroplasties - A comparison between the Harris hip score and the Oxford Hip Score. *J Arthroplasty*. 2005;20:1037–41.
- Gosens T, Hoefnagels NHM, de Vet RW, Dierckx WJA, van Langelaan EJ, Bultma SK, et al. The "Oxford Hip Score". *Acta Orthop*. 2005;76:204–11.
- Delunay C, Epinette JA, Dawson J, Murray D, Jolles BM. Cross-cultural adaptations of the Oxford-12 HP score to the French speaking population. *Orthop Traumatol Surg Res*. 2009;95:89–99.
- Naal FD, Seiverding M, Impellizzeri FM, von Knorck F, Mannion AF, Leunig M. Reliability and validity of the cross-culturally adapted German Oxford Hip Score. *Clin Orthop Relat Res*. 2009;467:952–7.
- Martinelli N, Longo UG, Marinuzzi A, Franceschini E, Costa V, Denaro V. Cross-cultural adaptation and validation with reliability, validity, and responsiveness of the Italian version of the Oxford Hip Score in patients with hip osteoarthritis. *Qual Life Res*. 2011;20:923–9.
- Paulsen A, Odgaard A, Overgaard S. Translation, cross-cultural adaptation and validation of the Danish version of the Oxford Hip Score

- assessed against generic and disease-specific questionnaires. *Bone Jt Res.* 2012;1:225–33.
25. Tuğay BU, Tuğay N, Güneş H, Hazar Z, Yüksel İ, Atila B. Cross-cultural adaptation and validation of the Turkish version of Oxford Hip Score. *Arch Orthop Trauma Surg.* 2015;135:879–89.
26. Lee Y-K, Chung CY, Park MS, Lee KM, Koo K-H, Lee DJ, et al. Transcultural adaptation and testing of psychometric properties of the Korean version of the Oxford Hip Score. *J Orthop Sci.* 2012;17:377–81.
27. Zheng W, Li J, Zhao J, Liu D, Xu W. Development of a valid simplified Chinese version of the Oxford Hip Score in patients with hip osteoarthritis. *Clin Orthop Relat Res.* 2014;472:1545–51.
28. Uesugi Y, Makimoto K, Fujita K, Nishii T, Sakai T, Sugano N. Validity and responsiveness of the Oxford Hip Score in a prospective study with Japanese total hip arthroplasty patients. *J Orthop Sci.* 2009;14:35–9. Elsevier.
29. Martínez JP, Arango AS, Castro AM, Martínez Rondanelli A. Validación de la versión en español de las escalas de Oxford para rodilla y cadera. *Rev Colomb Ortop y Traumatol.* 2016;30(2):61–6.
30. Altman RD. Criteria for classification of clinical osteoarthritis. *J Rheumatol Suppl.* 1991;27:10–2.
31. Wolf EJ, Harrington KM, Clark SL, Miller MW. Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. *Educ Psychol Meas.* 2013;73:913–34.
32. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med.* 2002;21:1331–5. John Wiley & Sons, Ltd.
33. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40:373–83.
34. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5 L). *Qual Life Res.* 2011;20:1727–36.
35. Ramos-Goñi J, Craig B, Oppe M, Ramallo-Fariña Y, Pinto-Prades J, Luo N, et al. How to handle data quality issues in EQ-5D-5 L valuation studies. *The Spanish case.* *Value Heal.* 2016;19A:376.
36. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, et al. The use of the Oxford hip and knee scores. *J Bone Jt Surg.* 2007;89:1010–4.
37. 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;60:34–42.
38. Beavers AS, Lounsbury JW, Richards JK, Huck SW, Skollis GJ, Esquivel SL. Practical considerations for using exploratory factor analysis in educational research. *Pract Assessment, Res Eval.* 2013;18:1–13.
39. Batista-Foguet JM, Coenders G, Alonso J. Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud. *Med Clin (Barc).* 2004;122:21–7.
40. Mulaik SA, James LR, Van Alstine J, Bennett N, Lind S, Sittwell CD. Evaluation of goodness-of-fit indices for structural equation models. *Psychol Bull.* 1989;105:430. American Psychological Association.
41. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: a review. *J Educ Res.* 2006;99:233–38. Taylor & Francis.
42. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16:297–334.
43. Cronin JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol.* 1993;78:98–104.
44. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics.* 1977;33(2):363–74.
45. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care.* 1989;27:5178–89. LWW.
46. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol.* 2003;56:395–407.
47. Schmitz JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol.* 2004;57:1008–18. Elsevier.
48. Weir JP. The intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005;19:231–40.
49. De Boer MR, De Vet HCW, Terwee CB, Moll AC, Völkers-Dieben HJM, Van Rens GHMB. Changes to the subscales of two vision-related quality of life questionnaires are proposed. *J Clin Epidemiol.* 2005;58:1260–8.
50. Impellizzeri FM, Mannoni AF, Nael FD, Leung M. Validity, reproducibility, and responsiveness of the Oxford Hip Score in patients undergoing surgery for femoroacetabular impingement. *Arthroscopy.* 2015;31:42–50. Elsevier.
51. Ozendoff M, Van Sael HF, Buskens E, Schijvers AJP, Marting LN, Verbout AJ, et al. Patient-reported outcome in total hip replacement. *J Bone Jt Surg.* 2004;86:801–8.
52. Norquist JM, Fitzpatrick R, Dawson J, Jenkinson C. Comparing alternative Rasch-based methods vs raw scores in measuring change in health. *Med Care.* 2004;42:21–25. LWW.
53. Harris KK, Price AJ, Beard DJ, Fitzpatrick R, Jenkinson C, Dawson J. Can pain and function be distinguished in the Oxford Hip Score in a meaningful way? an exploratory and confirmatory factor analysis. *Bone Joint Res.* 2014;3:305–9.
54. Paulsen A. Patient reported outcomes in hip arthroplasty registries. *Dan Med J.* 2014;61:8485.
55. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Jt Surg.* 1996;78-B:185–90.
56. Dawson J, Fitzpatrick R, Frost S, Gundle R, McLardy-Smith P, Murray D. Evidence for the validity of a patient-based instrument for assessment of outcome after revision hip replacement. *J Bone Joint Surg (Br).* 2001;83:1125–9.
57. Beard DJ, Harris K, Dawson J, Dol H, Murray DW, Carr AJ, et al. Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. *J Clin Epidemiol.* 2015;68:73–9. Elsevier Inc.
58. Norman GR, Sloan JA, Wywich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care.* 2003;41:582–92.
59. Strand U, Ljunggren AE, Bogen B, Ask T, Johnsen TB. The short-form McGill pain questionnaire as an outcome measure: test-retest reliability and responsiveness to change. *Eur J Pain.* 2008;12:917–25.
60. Thorborg K, Roos E, Bartels E, Petersen J, Holmich P. Validity, reliability and responsiveness of patient-reported outcome questionnaires when assessing hip and groin disability: a systematic review. *Br J Sports Med.* 2010;44:1186–96.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit



RESEARCH

Open Access



Validation of the Spanish version of the Oxford knee score and assessment of its utility to characterize quality of life of patients suffering from knee osteoarthritis: a multicentric study

Jesús Martín-Fernández^{1,2,3*}, Roberto García-Maroto⁴, Fco Javier Sánchez-Jiménez⁵, Alonso Bau-González⁶, Homero Valencia-García⁷, Blanca Gutiérrez-Teira⁸, Juan Carlos Arenaza^{9,3}, Lidia García-Pérez^{10,3}, Renata Linertová^{10,3} and Amaia Bilbao^{11,3}

Abstract

Background: Knee osteoarthritis (OA) represents a heavy burden for patients and the society as a whole. The Oxford Knee Score (OKS) is a well known tool to assess the quality of life in patients with Knee OA. The purpose of this study was to analyze the psychometric properties of the Spanish version of the OKS, including its reliability, validity, and responsiveness.

Methods: Prospective observational study that included 397 patients diagnosed with knee OA according to the criterion of the American Rheumatism Association, which were recruited in 3 different Spanish regions. Their self-perceived health-related quality of life (HRQL) was assessed through 3 questionnaires: a generic one (the EQ-5D-5 L) and two specific ones adapted to Spanish (the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) and the Oxford Knee Score (OKS)). The follow-up period was 6 months, and the acceptability of the OKS was evaluated, together with its psychometric properties, presence of ceiling and floor effects, validity, reliability, and sensitivity to change.

Results: The OKS was fully answered in 99.5% of cases, with no evidence of ceiling or floor effects. Its factor structure can be explained in a single dimension. Its discriminating capacity was very good compared to the groups generated by the WOMAC and the EQ-5D-5 L. The correlation of the OKS with the dimensions of the latter questionnaires was around 0.7. The test-retest reliability was excellent (ICC 0.993; CI 95%: 0.990–0.995) and so was its internal consistency (Cronbach's $\alpha = 0.920$). The effect size was 0.7 for moderate improvements in the HRQL, which is similar to that of the dimensions of the WOMAC and greater than for the EQ-5D-5 L. The minimum clinically significant difference that was detected by the questionnaire was 6.1 points, and the minimum detectable change was 4.4 points.

(Continued on next page)

* Correspondence: jmferrandez@salud.madrid.org

¹C³ Villamanta (CS, Navalcamero). Gerencia Asistencial de Atención Primaria. Servicio Madrileño de Salud, Avda Libertad 21 s/n, Villamanta, 28610 Madrid, Spain

²Facultad de Ciencias de la Salud. Universidad Rey Juan Carlos, Avda Atenas s/n, 28922 Alcorcón, Madrid, Spain

Full list of author information is available at the end of the article



© The Author(s). 2017 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

(Continued from previous page)

Conclusions: The Spanish-adapted version of the OKS is a useful, valid tool for assessing the perceived HRQL in patients suffering from knee OA, with psychometric properties similar to the WOMAC, and that allows for discriminating the patient's condition at a particular moment as well as for appraising changes over time.

Keywords: Quality of life, Health status, Osteoarthritis, knee, Validation studies

Background

Osteoarthritis (OA) is the most frequent joint disease, characterized by progressive articular cartilage loss that results in joint pain and functional impairment, which impacts the ability to perform daily-life activities. The prevalence of this type of disease is very high affecting 4% of the general population worldwide based on radiological diagnosis, and up to 20% in the case of specific population groups, such as women over 60 years [1]. Knee OA is a heavy burden for patients and the society as a whole. International studies have estimated that knee and hip OA constituted 0.7% of all disability adjusted life years (DALY) lost in 2010, a 40% increase with respect to 1990 [1]. Eighty-three percent of DALY lost due to OA are due to OA of the knee [2].

Knee OA entails a substantial impact on health related quality of life (HRQL) [3, 4]. HRQL is generally considered to incorporate the evaluation of functioning status as well as the patient's perception of their emotional functioning and social role. Since patients' responses vary greatly in the face of identical stressors, such as pain, HRQL is a crucial outcome measure [5]. The dimensions of HRQL most affected by knee OA are those related to physical activity and self-efficiency [6]. It seems that knee OA has a greater impact on the physical aspects of HRQL in the case of women, whereas men report worse scores on psychological-related scales [7]. Besides, HRQL predicts future inpatient and outpatient health care utilization and mortality in patients diagnosed of OA [8]. Therefore, measures of HRQL are important not only for assessing the burden of the disease or the results of any intervention, but also for helping informed decision-making in the allocation of often limited health resources [4].

In the case of knee OA, there are several specific tools to measure HRQL, some of which have been adapted and validated for the Spanish setting, such as the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) –a useful questionnaire for the assessment of OA of the lower limb [9, 10]–, the “Knee Society Clinical Rating System” (KSS) [11, 12], or the Knee Injury and Osteoarthritis Outcome Score (KOOS) [13]. Other questionnaires like the “Oxford Knee Score” (OKS) have not had their psychometric properties validated for our setting. The OKS is a brief, 12-item,

self-reported scale developed to measure the impact of total knee replacement surgery on the perception of HRQL by patients [14], and its scores and outcome interpretations have been slightly modified throughout the years it has been in use [15]. It is reported to be amongst the most sensitive, responsive, reliable, and valid patient-reported questionnaire for knee conditions [16]. It has been adapted and validated into Italian [17], Dutch [18], Chinese and Singapore English [19], German [20], French [21], Japanese [22], Portuguese [23], Korean [24], Persian [25], Greek [26], Spanish in Colombia [27], Arabic language [28], and Finnish language [29]. Owing to its good psychometric properties, it has been favorably compared to other widely used tools in different languages that are more difficult to administer [30, 31]. Although the OKS has been adapted to Spanish for Spain, its psychometric properties have not been assessed in the Spanish population setting. As far as we know, only the Dutch and Finnish language adaptations of OKS have been validated in a prospective manner, similarly to the original work by Dawson et al. [18, 29], whereas its factor structure has not been confirmed in any of its adapted versions.

In Spain, knee OA implies an enormous burden of illness for the people who suffer from it and for the whole society [32], and is therefore worthy of being measured. There are new instruments that serve this purpose, such as the OKS; however, for a questionnaire to be useful in culturally different areas with different languages, it must not only be translated into the new language but also adapted to account for any different or new cultural characteristics. The adaption must then be validated as the original version was. This work tackles the study of the psychometric properties of the OKS in its Spanish-adapted version, including its reliability, validity, and responsiveness.

Methods

Design

Prospective observational study. A population sample was recruited and followed up after 6 months.

Sampling and sample size

Opportunistic sampling of patients diagnosed with knee OA was performed both in traumatology and primary care consultations in Bizkaia, Madrid, and Tenerife.

Patients were included in a consecutive way between January and December 2015. All patients were chronic, and the knee OA was diagnosed according to the American Rheumatism Association's criterion [33], either by the clinician that included the patient in the study or from what was already recorded in the clinical history. Patients with OA from other regions and those suffering other comorbidities were also included. Subjects that did not properly understand or read Spanish and those diagnosed with any cognitive impairment were excluded.

The confirmatory factor analysis (CFA) set the minimum requirements to calculate sample size since it was the most stringent of the employed analytic methods in this regard. It was estimated that 300 patients would be needed, using a questionnaire with a single factor comprised of 12 items [34]. This sample size would also allow for estimating intraclass correlation coefficients (ICC) of >0.8 with precision values $<10\%$ [35].

All included patients provided written consent for participation and the study was approved by the relevant Ethics Committees for Clinical Research.

Variables

The personal variables recorded for each patient were age, gender, body mass index (BMI), joints affected by arthritis, previous joint replacement surgeries, and Charlson's index, which was calculated to assess comorbidity situations [36]. Patients answered three questionnaires, all in their Spanish version, in order to appraise their HRQL: a generic one (EQ-5D-5 L) [37], and two specific to OA (the WOMAC [9] and the OKS [14]).

The EQ-5D-5 L Spanish for Spain version has shown initial content and face validity [37]. This new version improves the old EQ-5D-3 L version, which had high internal consistency and reliability levels but, on the contrary, showed ceiling effect and low responsiveness [38]. The EQ-5D-5 L asks about current self-perception of health and is comprised of two parts. The first part includes 5 questions on general health: mobility, self-care, performance of daily-life activities, pain/discomfort, and anxiety/depression. Each dimension is measured on a scale from 1 to 5. A single weighted score for health condition is then obtained from these 5 questions, the so-called utility index, and the higher the score the better the health status [39]. The second part consists of a visual analogue scale (VAS) that ranges from 0 (worst health condition) to 100 (best health condition).

The WOMAC [9] is a self-administered questionnaire, specific to patients suffering from OA of the hip or knee. It has a multidimensional scale comprised of 24 items clustered according to 3 domains: pain (5 items), stiffness (2 items), and physical functionality (17 items). Its Likert version, where each item receives a score from 0 to 4 corresponding to the different intensity levels of

the response (none, light, moderate, severe, extreme), was chosen. This score is summed and standardized from 0 (best ability) to 100 (worst ability). The greater the score, the worse the health condition of the patient. This questionnaire has been adapted and validated for our setting. The adapted version of the WOMAC questionnaire showed high convergent validity, internal consistency (Cronbach's α ranging from 0.81 to 0.93), and test-retest reliability. The responsiveness test showed effect sizes ranging from 1.5 to 2.2 in patients that had undergone hip replacement [10].

The OKS is a self-administered questionnaire that can be answered via "face to face" interviews or mailed-in by the patient after completion. It contains 12 questions, with 5 possible answers each, intended to evaluate the patient's perception of quality of life over the last 4 weeks. It has been used both to assess the baseline situation and to study changes after prosthetic implants in patients suffering from knee OA. Each answer is given a score from 0 to 4, where 4 is the best possible result [15]. After being summed up, a total score is obtained that ranges from 0 to 48, where 48 is the best possible outcome. The Spanish-adapted version was created under agreement with the Oxford University Innovation, following a process of translation and inverse retro-translation (Additional file 1).

Subjects recruited in Madrid were interviewed 7 to 15 days after the inclusion visit, and the OKS questionnaire was repeated after ensuring that there were no changes in their health condition. All included patients were interviewed again after a follow-up period of 6 months: they were asked if they had undergone replacement surgery, the EQ-5D-5 L, WOMAC, and OKS questionnaires were repeated, and transition questions were posed to assess if their general health self-perception had suffered any changes.

Statistical analysis

Continuous variables were described by their measures of central tendency and dispersion, whereas discrete variables were described by their percentages. Confidence intervals were set at 95%.

Acceptability and floor and ceiling effects

The number of unfilled questionnaires and unanswered questions was noted.

Ceiling or floor effects were considered to be present if more than 15% of respondents reported the highest or lowest possible score, respectively [40].

Analysis of the psychometric properties

Validity

The validity of the construct was appraised via an explanatory factor analysis (EFA) that analyses the unidimensionality

of the questionnaire. The sampling adequacy was checked using Bartlett's test of sphericity and Kaiser-Meyer-Olkin (KMO) test. The null hypothesis of Bartlett's test is that the correlations matrix is a singular matrix. The rejection of this hypothesis allows for confirming the existence of linear relationships between the factors and the explained variable. The KMO sampling adequacy test is a measure of the covariance among variables and values >0.90 are considered to be optimal [41]. Both factor loading (values >0.40 were considered optimal) and communalities were noted, which together account for the percentage of the item's variance explained by each factor.

To complement our results, a CFA for categorical variables was also performed. The robust unweighted least squares estimator was used and several fit indices were calculated [42, 43]: the root mean square error of approximation (RMSEA), for which a value <0.08 was considered acceptable, and the Tucker-Lewis Index (TLI) and Comparative Fit Index (CFI), both of which had to be >0.95 to be considered satisfactory [44]. Factor loadings were also examined and those ≥ 0.40 were considered acceptable. Therefore, the model was considered adequate when these acceptability criteria were met.

The scores obtained through the OKS were compared to the terciles of the distributions obtained from the EQ-5D-5 L and WOMAC questionnaires in order to assess the validity of the known groups.

Convergent validity was checked through the correlations of the OKS scale with the WOMAC and EQ-5D-5 L (utility index and VAS) scales. Pearson's r or Spearman's ρ were used to study such correlations, and 0.7 was set as the threshold for considering strong associations to be present [40].

Reliability

Internal consistency was tested using Cronbach's α [45] that was obtained from the scores of the inclusion visit. This coefficient summarizes internal correlations of all the elements of a scale. The greater the coefficient (range 0.0–1.0), the greater the internal consistency of the scale and the greater the probability for a single dimension to be underlying the questionnaire. For a single-dimension tool comprising 12 components, Cronbach's α is expected to reach values >0.85 in order for its internal consistency to be considered optimal [46].

The test-retest reliability was studied in the subsample from Madrid, and ICCs were used to compare the test against the retest scores. According to the classification proposed by other reliability measures [47], ICC values >0.7 are considered acceptable and >0.9 optimal.

Responsiveness

The OKS questionnaire was repeated at a follow-up period of 6 months to evaluate its responsiveness to

changes resulting from disease progression. In order to assess changes in the knee condition compared to the 6 previous months as perceived by patients, transition questions were posed and answered on a scale comprising 5 answers (much worse, slightly worse, same, slightly better, or much better than before). These questions were aimed at appraising the sensitivity of the OKS questionnaire to change. In the case of the WOMAC, transition questions were answered on the same scale, but they were specific for each of its domains (pain, stiffness, and physical functionality).

Changes were appraised with the OKS by subtracting initial from final scores, so that positive values indicate an improvement in general condition. The procedure was the same with the EQ-5D-5 L but, in the case of the WOMAC, the final scores were subtracted from the initial ones so that positive values also indicated improvements. Transition questions were posed to every group of patients in order to see if significant changes had occurred, and basal scores were contrasted against those at 6 months of follow-up. The relationship between the median and standard deviation was calculated to determine the effect size of the change for each group of patients: values >0.5 were regarded as moderate change, and values >0.8 as large change [48]. The effect size was then compared to the one obtained from the WOMAC and EQ-5D-5 L scales.

Furthermore, the minimal clinically important difference (MCID) and the minimal detectable change (MDC) were determined. These two measures are related to responsiveness, but are more clinically oriented and focused at the individual level. Average change in patients that had experienced moderate improvement in their condition (reported feeling "slightly better") was used to calculate MCID at the 6 months follow-up [49].

The MDC expresses the minimal magnitude of change above which the observed change is likely to be real and not just measurement error. For estimation of MDC, the standard error of measurement (SEM) was determined, which quantifies the precision of individual scores on a test. The SEM was estimated as the square root of the mean square error term from the ANOVA [50, 51]. From the SEM, the MDC was derived as follows [40, 50]: $MDC = SEM \times z\text{-score} \times \sqrt{2}$. A 95% confidence level ($MDC_{95\%}$) was set, corresponding to a z -value of 1.96. The interpretation of $MDC_{95\%}$ is that if a patient shows a score change equal to or greater than the $MDC_{95\%}$ threshold, it is possible to state with 95% confidence that this change is reliable and not the result of a measurement error. Finally, the MCID was divided by the $MDC_{95\%}$ to determine if the MCID surpassed the $MDC_{95\%}$ [52]: if this ratio exceeded 1, the MCID could be discriminated from measurement error.

All effects were considered statistically significant at $p < 0.05$. The statistical analyses were performed using SPSS 18.0 and Mplus 6.1 software.

Results

A total of 397 patients were included: 158 in Bizkaia, 158 in Madrid, and 81 in Tenerife. Of them, 36.8% were recruited at primary care, 55.2% at traumatology, and 8.0% at rheumatology consultations. The mean time elapsed since diagnosis was 61.6 months (CI 95%: 55.6–67.6 months). Women comprised 69.8% (CI 95%: 65.3–74.3%) of the sample, with an average age of 71.4 years (CI 95%: 70.5–72.3 years).

In terms of the knee affected by OA, in 27.7% (CI 95%: 23.3–32.1%) of cases it was the right knee, in 30.0% (CI 95%: 25.5–34.5%) the left knee, and in 43.3% (CI 95%: 38.4–48.2%) both knees.

Total knee replacement surgery had been previously performed in 18.1% (CI 95%: 14.3–21.9%) of cases. The average Charlson's comorbidity index was 0.8 points (CI 95%: 0.7–0.9), and mean BMI was 29.7 (IC 95%: 29.2–30.2).

Table 1 shows the outcome from the responses given by patients to the OKS, WOMAC, and EQ-5D-5 L questionnaires.

Acceptability and floor and ceiling effects

Information was obtained in 395 cases (99.5%; CI 95%: 98.8–100%) which allowed summarizing the results from the OKS questionnaire. Questions 7, 9, and 12 were answered in all cases, and questions 1, 2, 3, 5, 6, 10, and 11 in all cases but one. Question 8 was not answered in 2 occasions, and question 4 in 6 cases. All possible answers, namely all ranges of response (0 to 4), were posed for every question. Only in the case of question 7, 48% of responses were concentrated in the top score. Only in

the case of questions 1 and 7, less than 10% of responses clustered into the bottom end of the scale (0 and 1), which did not happen for the top end in any case. For the total score, there was no aggregation at the low end of the scale and only 0.25% and 0.61% of the responses scored 48 out of 48 possible points in the inclusion visit or in the six month visit respectively. Hence, the presence of floor or ceiling effects was ruled out.

Validity

With regards to the validity of the construct, a unidimensional structure was found in the EFA with a single factor that explained 55.5% of variance (KMO = 0.946, Bartlett's test of sphericity $\chi^2 = 2597$, 66 degrees of freedom, $p < 0.001$). All factor loadings were >0.50 , and commonalities were >0.40 except for questions 4 and 8 (Table 2).

The results of the CFA (Fig. 1) showed excellent fit indices: (a) the RMSEA was 0.076, that is <0.08 ; and (b) the CFI and TLI were 0.981 and 0.977, respectively, exceeding the benchmark of 0.95. Factor loadings were all statistically significant ($P < 0.001$), ranging from 0.58 to 0.86 (Fig. 1).

The validity of known groups, which measures the discriminatory capacity of the questionnaire, can be checked in Table 3, where mean scores and 95% CI of the OKS are shown for the various terciles of the WOMAC and EQ-5D-5 L scales distribution. Differences between the three groups are clearly observed in the OKS scores, with average changes of 5.6 and 11.9 points per tercile.

Table 4 shows the correlations between the OKS scores and WOMAC domains or EQ-5D-5 L VAS and utilities. Due to the different ways in which scores are presented on the scales, negative correlations with the

Table 1 Summary of the outcome from the OKS, WOMAC, and EQ-5D-5 L questionnaires

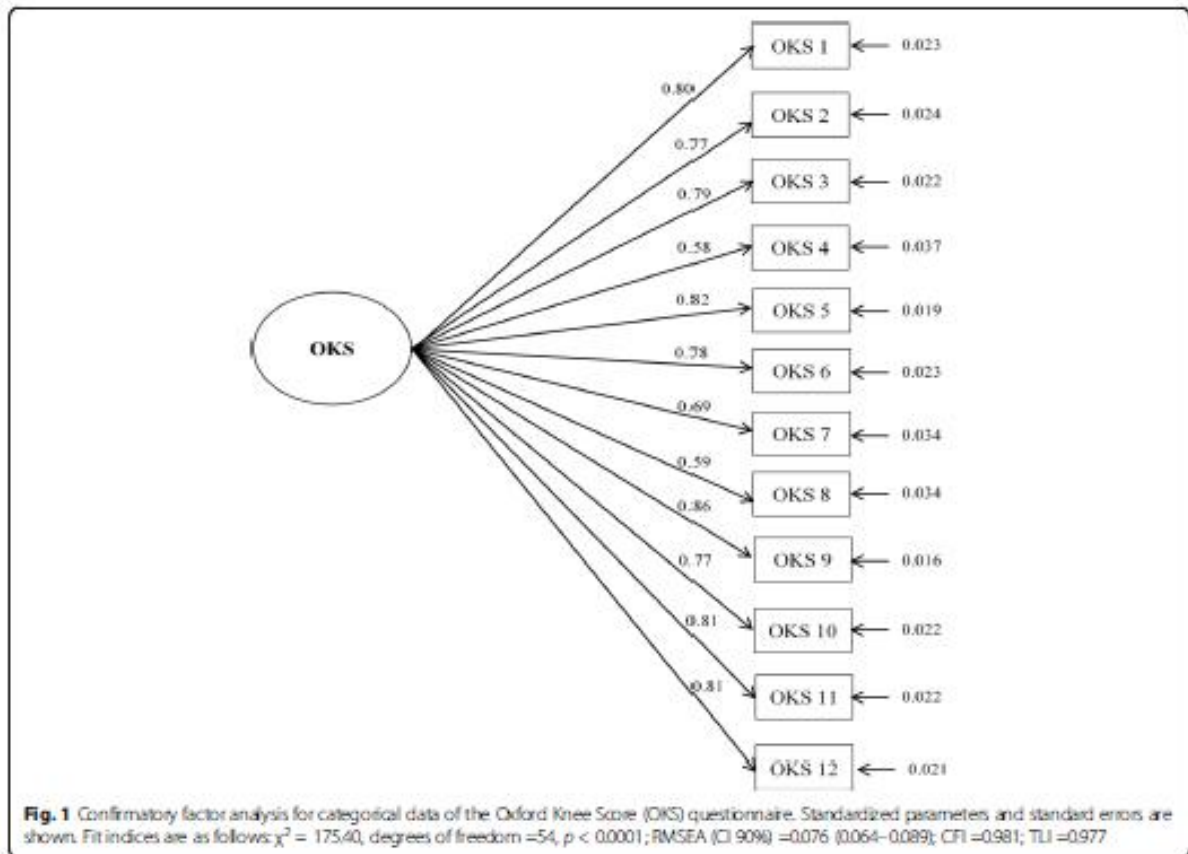
	n	Average score (IC 95%)	Median (Interquartile range)
OKS	395	22.0 (21.0–22.9)	20.0 (14.0–27.0)
WOMAC pain	396	46.2 (44.1–48.3)	44.4 (30.0–60.0)
WOMAC rigidity	396	45.5 (42.9–48.0)	46.4 (25.0–62.5)
WOMAC impairment	397	52.7 (50.5–55.0)	48.5 (25.0–62.5)
EQ-5D-5 L utility	393	0.54 (0.52–0.57)	0.62 (0.39–0.74)
EQ-5D-5 L VAS	390	57.0 (55.2–59.5)	60.0 (45.0–75.0)

OKS Oxford Knee Score. Range 1–48. The better the score, the better the health condition
WOMAC Western Ontario and McMaster Universities Osteoarthritis Index. Scale range 1–100. The better the score, the worse the health condition
EQ-5D-5 L Utility score range 0–1, where 0 = condition comparable to death, and 1 = perfect health, although negative scores are allowed
EQ-5D-5 L VAS: Visual Analogue Scale. Range 0–100, where 0 = worst health condition, and 100 = best conceivable health condition

Table 2 Exploratory factor analysis of items in the Oxford Knee Score (OKS)

Variable	EFA 1 factor	
	Factor loading	Commonality
OKS 1	0.761	0.579
OKS 2	0.751	0.564
OKS 3	0.774	0.599
OKS 4	0.572	0.327
OKS 5	0.796	0.634
OKS 6	0.773	0.597
OKS 7	0.670	0.449
OKS 8	0.605	0.365
OKS 9	0.839	0.705
OKS 10	0.767	0.589
OKS 11	0.788	0.620
OKS 12	0.794	0.630

EFA Exploratory factor analysis
OKS Oxford Knee Score



WOMAC and positive ones with the EQ-5D-5 L were to be expected. All associations were strong except for rigidity on the WOMAC scale, whose correlation was at the limit of the required threshold, and the EQ-5D-5 L VAS.

Reliability

Regarding internal consistency, Cronbach's α was 0.920 for the OKS questionnaire.

The ICC for the 158 subjects that repeated the questionnaire at 7 and 14 days after the inclusion visit was

0.993 (CI 95%: 0.990–0.995) and Cronbach's α was 0.997 at both check points.

Responsiveness

After 6 months, follow-up was possible in the case of 331 subjects. Of those, 42 people had undergone joint replacement surgery. Thirty-three patients (10.1%; CI95%: 6.7–13.2%) received some sort of rehabilitation or physical therapy during this period. One hundred and one patients (30.5%; CI 95%: 25.6–35.5%) reported feeling "slightly better" or "much better" when asked about

Table 3 Average scores of the OKS in patients after being clustered according to the terciles obtained from the WOMAC and EQ-5D-5 L questionnaires

	OKS score for the lower tercile of the questionnaire distribution (CI 95%)	OKS score for the middle tercile of the questionnaire distribution (CI 95%)	OKS score for the top tercile of the questionnaire distribution (CI 95%)
WOMAC pain	31.4 (30.0–32.7)	20.5 (19.5–21.6)	13.8 (12.7–14.9)
WOMAC rigidity	29.7 (28.2–31.3)	22.1 (20.8–23.4)	14.4 (13.3–15.4)
WOMAC impairment	32.0 (30.8–33.3)	20.1 (19.1–21.0)	13.1 (12.1–14.0)
EQ-5D-5 L utility	13.6 (12.5–14.6)	21.1 (20.0–22.2)	31.1 (29.7–32.4)
EQ-5D-5 L VAS	15.7 (14.0–17.3)	21.5 (20.3–22.7)	27.1 (25.3–28.9)

OKS Oxford Knee Score
WOMAC Western Ontario and McMaster Universities Osteoarthritis Index
VAS Visual Analogue Scale

Table 4 Correlations (Pearson's *r*) between the scores from the OKS, WOMAC scales, and EQ-5D-5 L scales (utility index and VAS)

	OKS	WOMAC pain	WOMAC rigidity	WOMAC impairment	EQ-5D-5 L Utility	EQ-5D-5 L VAS
OKS	—					
WOMAC pain	-0.745*	—				
WOMAC rigidity	-0.641*	0.643*	—			
WOMAC impairment	-0.849*	0.799*	0.719*	—		
EQ-5D-5 L utility	0.756*	-0.623*	-0.504*	-0.706*	—	
EQ-5D-5 L VAS	0.530*	-0.472*	-0.361*	-0.516*	0.547*	—

**p* < 0.001

OKS Oxford Knee Score

WOMAC Western Ontario and McMaster Universities Osteoarthritis Index

Higher scores indicate better health condition in the OKS and EQ-5D-5 L, and opposite in the case of the WOMAC

the knee that caused their inclusion in the study, and 143 (43.2%; CI 95%: 37.9–48.5%) stated they felt "slightly worse" or "much worse".

Tables 5 and 6 show the average change in the scores of the diverse questionnaires employed when the patient had perceived a change in health condition. When the OKS was used, the effect size of the change was 0.69 for subjects that stated feeling "slightly better" and 1.60 if they felt "much better". The effect size was lower for negative changes, with a value of 0.24 for moderate ("slightly worse") changes and 0.57 in the case of substantial ("much worse") negative changes. There was a clear gradient in the score depending on the change perceived by the patient, which is significantly different for groups that reported feeling "slightly worse", "slightly better", or "much better". There is a small overlap between those who felt "much worse" and "slightly worse". The tool proved to be more sensitive than the EQ-5D-5 L and worked in a similar way to the WOMAC scales on pain and impairment, whereas the scale on rigidity was less sensitive to change.

For subjects that experienced a "moderate" subjective improvement, the average change in the OKS was 6.1 points (SD = 8.9), which was used to estimate the MDCl. The SEM was estimated to be 1.5, so $MDC_{95\%}$ was calculated to be 4.38, which means that the ratio $MDCl / MDC_{95\%}$ was 1.4.

Discussion

The Spanish version of the OKS questionnaire is a reliable, sensitive to changes, valid tool to measure HRQL in patients that suffer from knee OA. Given the extraordinarily high response rate, it also is a well accepted questionnaire.

The validity of the OKS was assessed from different perspectives, although its "apparent" validity has not been tested since it is an adaptation. Discriminatory or known-groups validity seems adequate since the outcome score differs greatly in subjects with very different scores on the WOMAC or the EQ-5D-5 L scales. Additionally, it does not show significant ceiling or floor effects that compromise such discriminatory capacity, as has been previously pointed out in other adaptations [17, 20, 21, 29].

The convergent validity of the tool seemed appropriate. Correlations of the OKS adapted-version with the specific scales of the WOMAC or the generic scales of the EQ-5D-5 L were stronger than those found between the original version and other generic tools that measure HRQL [14, 30]. In the case of adaptations of the OKS to other languages, like Portuguese [23] or German [20], these correlations were similar or slightly stronger. This way of measuring convergent validity offered better results for the OKS than those reported in other questionnaires like the Spanish version of the KSS [12].

Table 5 Changes in the OKS, and EQ-5D-5 L questionnaires observed after a follow-up period of 6 months in patients that reported changes in their condition

	The knee condition is "much worse" <i>N</i> = 59		The knee condition is "slightly worse" <i>N</i> = 84		The knee condition is "slightly better" <i>N</i> = 56		The knee condition is "much better" <i>N</i> = 45	
	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.
OKS	-4.5 (-6.6– -2.4)	0.57	-1.6 (-2.9– -0.2)	0.24	6.1 (3.7–8.5)	0.69	17.4 (14.1–20.7)	1.60
EQ-5D-5 L utility	-0.17 (-0.25– -0.08)	0.55	-0.01 (-0.06–0.03)	0.07	0.14 (0.07–0.20)	0.43	0.30 (0.21–0.39)	0.96
EQ-5D-5 L VAS	-4.6 (-12.8–3.7)	0.15	-0.74 (-4.9–3.4)	0.04	7.9 (3.0–12.8)	0.58	13.2 (9.0–17.4)	0.94

E.S. Effect size

OKS Oxford Knee Score

Basal scores were subtracted from final scores, so that positive outcomes indicate improvement

Table 6 Changes in the WOMAC questionnaire observed after a follow-up period of 6 months in patients that reported changes in their condition

	Pain is "much worse" N = 51		Pain is "slightly worse" N = 82		Pain is "slightly better" N = 61		Pain is "much better" N = 47	
WOMAC pain	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.
	-9.9 (-14.9--4.8)	0.55	-2.0 (-5.2-1.3)	0.13	17.7 (13.1-22.3)	0.98	28.8 (21.5-36.1)	1.15
WOMAC rigidity	Rigidity is "much worse" N = 44		Rigidity is "slightly worse" N = 85		Rigidity is "slightly better" N = 52		Rigidity is "much better" N = 42	
	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.
	-13.1 (-20.9--5.3)	0.51	-0.7 (-5.9-4.5)	0.03	12.3 (5.1-19.4)	0.48	30.1 (21.9-38.2)	1.15
WOMAC impairment	Impairment is "much worse" N = 47		Impairment is "slightly worse" N = 90		Impairment is "slightly better" N = 47		Impairment is "much better" N = 31	
	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.	Average change (CI 95%)	E.S.
	-10.5 (-15.2--5.7)	0.65	-2.4 (-5.5-0.7)	0.16	18.4 (12.3-24.4)	0.89	38.6 (29.7-47.6)	1.58

E.S. Effect size

WOMAC Western Ontario and McMaster Universities Osteoarthritis Index

Final scores were subtracted from basal scores. This way, positive outcomes indicate improvement

Construct validity of the OKS was also studied. Its factorial structure sustains the unidimensionality of the questionnaire. In the EFA, all items were found to consistently saturate the same factor and showed higher values than the English version [14]. Acceptable values of RMSEA were obtained during the CFA, and TLI and CFI were optimal [44]. Although the possibility of disaggregating pain and impairment components from the OKS has been proposed, this unifactorial structure seems to be the most solid one [53], which is supported by this outcome.

Internal consistency was better than for the original scale in the inclusion period (Cronbach's $\alpha = 0.92$ vs. 0.87) [14]. The test-retest reliability was very high and the obtained values, measured through the ICC, allowed to qualify the tool as reliable [40].

The discriminatory capacity of the questionnaire was adequate, which accounts for its ability to distinguish between individuals in different situations, but the tool can also be used to study the perception changes of a single person's situation, which means that its evaluative capability is adequate [54]. In fact, this tool was designed for that purpose and the outcome of this study supports this type of use. The effect size of the change for "moderate" positive changes was similar to the WOMAC but slightly lower than the set benchmark of 0.8. In the validation assessment of the original version, the observed effect size of the change after surgery was 2.1 [14], which is only comparable to really substantial improvements (effect size = 1.6) since it was tested in patients that had undergone knee replacement. The evaluative capacity was greater to detect positive than negative changes, as is the case with other questionnaires [55], although even in the case of negative changes the capacity of this tool

is similar or higher than the WOMAC, and greater than the generic EQ-5D-5 L questionnaire.

The MCID was 6.1 in the case of subjects that had experienced moderate improvement. Values of MCID between 3 and 5 points had been proposed for the validation of the original version [15] and confirmed in subsequent studies [56], although these studies only included subjects that had undergone knee replacement surgery. The $MCD_{95\%}$ was 4.38, which represents the lowest score change (at the particular patient level) that is not the result of measurement error of the instrument, and can be understood as the lowest bound of real change, although it may not indicate clinical significance [50]. The ratio between the MCID and $MCD_{95\%}$ was higher than 1, indicating that the MCID can be discriminated clearly from measurement error.

This work has some limitations. The studied subjects may not be representative of the national population. Patients were included from different regions and at different stages of the disease evolution, although we did not record or classify the knee OA severity of each patient. Besides, there are intrinsic limitations to the methodology used to assess the psychometric characteristics (the classical test theory), with its assumptions and weaknesses, but the validation process has been complemented with a CFA specific for categorical data that scrutinizes such deductive assumptions using statistical analysis [42].

The outcome of this study allows for proposing the application of the OKS in those situations where the original version has been used, such as measuring HRQL improvement after knee replacement surgery [30] or studying surgery-related factors [57-59], but also to discriminate between patients in different clinical situations

and to appreciate their evolution with time in view of its capacity to detect “moderate” improvements in patients.

Conclusions

The Spanish adaptation of the OKS questionnaire is a valid tool for assessing the perception of HRQL of patients suffering from knee OA. It is well accepted by patients and shows psychometric properties that support its usefulness both for the assessment of a patient's condition and its subsequent evolution. Its comparative utility is quite similar to that of tools that have been extensively used after their adaptation, like the WOMAC questionnaire. The incorporation of this type of tools in usual clinical practice will allow for appraising, in a valid and reliable way, the patient's self-perception of HRQL as well as the outcome of health interventions addressed at them.

Additional file

Additional file 1: Spanish-adapted version of the Oxford Knee Score - Spanish (Spain). (DOCX 23 kb)

Abbreviations

BMI: Body Mass Index; CFA: Confirmatory Factor Analysis; CFI: Comparative Fit Index; CI95%: 95% Confidence Interval; EFA: Exploratory Factor Analysis; GDP: Gross Domestic Product; HRQL: Health-Related Quality of Life; ICC: Intraclass Correlation Coefficient; KMO test: Kaiser-Meyer-Olkin test; MCDI: Minimal Clinically Important Difference; MDC: Minimal Detectable Change; OA: osteoarthritis; OKS: Oxford Knee Score; RMSEA: Root Mean Square Error of Approximation; SEM: Standard Error of Measurement; TLI: Tucker-Lewis Index; VAS: Visual Analogue Scale; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index

Acknowledgements

To Oxford University Innovation for providing us the adapted version of the Oxford Knee Score - Spanish (Spain).

Funding

This study has been financed by the Instituto de Salud Carlos III and the FEDER (European Regional Development Fund) (PI1300560, PI1300518 y PI1300648).

Availability of data and materials

The datasets used or analyzed during the current study are available from the corresponding author upon reasonable request.

Authors' contributions

Conceived and designed the experiments: IMF, RGM, LGP, RL, AB. Performed the experiments: IMF, FJS, ABG, HVG BGT, JCA. Analyzed the data: IMF, AB. Wrote the paper: IMF, AB. Revised and approved the article: IMF, RGM, FJS, ABG, HVG BGT, JCA, LGP, RL, AB. All authors read and approved the final manuscript.

Ethics approval and consent to participate

All patients gave written consent for participation.

This study received the approval of the following Ethics Committees for Clinical Research: ECCR of Euskadi (PI2014050), Hospital Fundación Jiménez Díaz (PC 80/2013_HRIC), Hospital Universitario de Fuenlabrada (APR 14-27), Hospital Universitario Fundación Alcorcón (14/19), Hospital Universitario de Canarias (2014-109), Hospital Universitario Nuestra Señora de Candelaria (PI-09/15).

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹C/ Villamarta (CS Navalmorera). Gerencia Asistencial de Atención Primaria. Servicio Madrileño de Salud, Avda Libertad 21 s/n, Villamarta, 28610 Madrid, Spain. ²Facultad de Ciencias de la Salud, Universidad Rey Juan Carlos, Avda Atenas s/n, 28922 Alcorcón, Madrid, Spain. ³Red de Investigación en Servicios de Salud en Enfermedades Crónicas, Bilbao, Spain. ⁴Servicio de Traumatología. Hospital Universitario Clínico San Carlos. Servicio Madrileño de Salud, C/ Profesor Martín Lagos, S/N, 28040 Madrid, Spain. ⁵CS Gregorio Marañón. Gerencia Asistencial de Atención Primaria. Servicio Madrileño de Salud, Calle Polvoranca, 65. 28923 Alcorcón, Madrid, Spain. ⁶Servicio de Traumatología. Hospital Universitario Rey Juan Carlos, C/ Gladioso s/n, 28933 Móstoles, Madrid, Spain. ⁷Servicio de Traumatología. Hospital Universitario Fundación Alcorcón, C/ Budapest, 1 28922 - Alcorcón, Madrid, Spain. ⁸C. S. El Soto. Gerencia Asistencial de Atención Primaria. Servicio Madrileño de Salud, Avenida Olímpica, 38, 28935 Móstoles, (Bizkaia), Spain. ⁹Servicio de Traumatología y Cirugía Ortopédica, Hospital Universitario Basurto (Osakidetza), Avda. Montevideo, 18, 48013 Bilbao, (Bizkaia), Spain. ¹⁰Fundación Canaria de Investigación Sanitaria (FUNCANIS), Camino Candelaria, Nº 44, 1ª planta, 38109 El Rosario (Santa Cruz de Tenerife), Spain. ¹¹Unidad de Investigación, Hospital Universitario Basurto (Osakidetza), Avda. Montevideo, 18, 48013 Bilbao, (Bizkaia), Spain.

Received: 20 February 2017 Accepted: 19 September 2017

Published online: 29 September 2017

References

- Cross M, Smith E, Hoy D, Nolte S, Ackerman J, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis*. 2014;73:1323-30.
- Yos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380:2163-96.
- Alkan BM, Fidan F, Tosun A, Ardicoglu O. Quality of life and self-reported disability in patients with knee osteoarthritis. *Mod Rheumatol*. 2014;24:166-71.
- Kadali AA, Lamm CJ, de Vedier MG, Engstrom G, Turkiewicz A, Lohmander LS, et al. Association of knee pain and different definitions of knee osteoarthritis with health-related quality of life: a population-based cohort study in southern Sweden. *Health Qual. Life Outcomes. Health Qual Life Outcomes*. 2016;14:121.
- Farr J, Miller LE, Block JE. Quality of life in patients with knee osteoarthritis: a commentary on nonsurgical and surgical treatments. *Open Orthop J*. 2013;7:619-23.
- Sundén A, Bidahl C, Magnusson SP, Johnson B, Gyllenstein AL. Physical function and self-efficacy – important aspects of health-related quality of life in individuals with hip osteoarthritis. *Eur J Phys*. 2013;15:151-9.
- Fang W-H, Huang G-S, Chang H-F, Chen C-Y, Kang C-Y, Wang C-C, et al. Gender differences between WOMAC index scores, health-related quality of life and physical performance in an elderly Taiwanese population with knee osteoarthritis. *BMI Open*. 2015;5:e008542.
- Singh J, Nelson D, Rink H, Nichol K. Health-related quality of life predicts future health care utilization and mortality in veterans with self-reported physician-diagnosed arthritis: The veterans arthritis quality of life study. *Semin Arthritis Rheum*. 2005;34:755-66.
- Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol*. 1988;15:1833-40.

10. Escobar A, Quintana JM, Bilbao A, Adánzte J, Guenaga JJ. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. *Western Ontario and McMaster Universities Osteoarthritis Index. Clin Rheumatol.* 2002;21:466–71.
11. Irgang JJ, Anderson AF, Boland AL, Hamer CD, Kurosaki M, Neyret P, et al. Development and validation of the international knee documentation committee subjective knee form. *Am J Sports Med. American Orthopaedic Society for Sports Medicine.* 2001;29:600–13.
12. Ares O, Castellet E, Maculé F, León V, Montañez E, Freire A, et al. Translation and validation of "The Knee Society Clinical Rating System" into Spanish. *Knee Surgery, Sport, Traumatol. Arthrosc.* 2013;21:2618–24.
13. Vaquero J, Longo UG, Fortiol F, Martindill 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 Surgery, Sport, Traumatol. Arthrosc.* 2014;22:104–8.
14. Dawson J, Fitzpatrick R, Murray Q, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br.* 1998;80:63–9.
15. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr A, et al. The use of the Oxford hip and knee scores. *J Bone Jt Surg.* 2007;89:1010–4.
16. Garratt AM, Brazley S, Gillespie WJ. Patient-assessed health instruments for the knee: A structured review. *Rheumatology.* 2004;43:1414–23.
17. Padua R, Zanoli G, Ceccarelli E, Romanini E, Bondi R, Campi A. The Italian version of the Oxford 12-item Knee Questionnaire/cross-cultural adaptation and validation. *Int Orthop.* 2008;32:214–6.
18. Haverkamp D, Bruggem SIM, Serevett IN, Bankevoort L, van Dijk CN. Translation and validation of the Dutch version of the Oxford 12-item knee questionnaire for knee arthroplasty. *Acta Orthop.* 2005;76:347–52.
19. Xie F, Li S-C, Lo N-N, Yeo S-J, Yang K-Y, Yeo W, et al. Cross-cultural adaptation and validation of Singapore English and Chinese Versions of the Oxford Knee Score (OKS) in knee osteoarthritis patients undergoing total knee replacement. *Osteoarthritis Cartil.* 2007;15:1019–24.
20. Naal FD, Impellizzeri FM, Sieverding M, Lohli M, von Knoch F, Mannion AF, et al. The 12-item Oxford Knee Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with osteoarthritis of the knee. *Osteoarthritis Cartil Elsevier Ltd.* 2009;17:49–52.
21. Jenny J, Diesinger Y. Validation of a French version of the Oxford knee questionnaire. *Orthop Traumatol Surg Res Elsevier Masson SAS.* 2011;97:267–71.
22. Takeuchi R, Sawaguchi T, Nakamura N, Ishikawa H, Saito T, Goldhahn S. Cross-cultural adaptation and validation of the Oxford 12-item knee score in Japanese. *Arch Orthop Trauma Surg.* 2011;131:347–54.
23. Gonçalves RS, Tomás AM, Martins DI. Cross-cultural adaptation and validation of the Portuguese version of the Oxford Knee Score (OKS). *Knee. Elsevier B.V.* 2012;19:344–7.
24. Eun IS, Kim OG, Kim CK, Lee HS, Lee JS. Validation of the Korean Version of the Oxford Knee Score in Patients Undergoing Total Knee Arthroplasty. *Clin Orthop Relat Res.* 2013;471:600–5.
25. Ebrahimi-zadeh MH, Makhmaltaf H, Brjandinejad A, Soltani-Moghaddeh SH. Cross-cultural adaptation and validation of the Persian version of the Oxford knee score in patients with knee osteoarthritis. *Iran J Med Sci.* 2014;39:529–35.
26. Strimpakos N, Dapkin F, Papsachristou A, Kapreli E. The 12-item Oxford knee score: cross-cultural adaptation into Greek and assessment of its psychometric properties. *Physiotherapy The Chartered Society of Physiotherapy.* 2015;101:1445–6.
27. Martínez JP, Arango AS, Castro AM, Martínez RA. Validación de la versión en español de las escalas de Oxford para rodilla y cadera. *Rev Colomb Ortop y Traumatol.* 2016;30:61–6.
28. Alghadir AH, Al-Esa ES, Anwer S. Cross-cultural adaptation and psychometric analysis of the Arabic version of the Oxford knee score in adult male with knee osteoarthritis. *BMC Musculoskelet Disord.* 2017;18:190.
29. Reito A, Jänistö A, Birnén E, Sleytö E, Remes V, Huhtala H, et al. Translation and validation of the 12-item Oxford knee score for use in Finland. *BMC Musculoskelet Disord.* 2017;18:74.
30. Benson T, Williams DH, Rotts HW. Performance of EQ-5D, howRu and Oxford hip & knee scores in assessing the outcome of hip and knee replacements. *BMC Health Serv Res BMC Health Serv Res.* 2016;16:512.
31. Goldhahn S, Takeuchi R, Nakamura N, Nakamura R, Sawaguchi T. Responsiveness of the Knee Injury and Osteoarthritis Outcome Score (KOOS) and the Oxford Knee Score (OKS) in Japanese patients with high tibial osteotomy. *Sci J Orthop.* 2017.
32. Loza E, Lopez-Gomez JM, Abasolo L, Maese JJ, Carrona L, Batlle-Gualda E. Economic burden of knee and hip osteoarthritis in Spain. *Arthritis Rheum.* 2009;61:158–65.
33. Altman R, Asch E, Bloch D, Boie 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;29:1039–49.
34. Wolf EJ, Hamington KM, Clark SL, Miller MW. Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Property. *Educ Psychol Meas.* 2013;73:913–34.
35. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat. Med. John Wiley & Sons, Ltd.* 2002;21:1331–5.
36. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1967;40:373–83.
37. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011;20:1727–36.
38. Badia X, Roset M, Montserrat S, Herdman M, Segura A. (The Spanish version of EuroQol: a description and its applications. European Quality of Life scale). *Med. Clin. (Barc).* 1999;112 Suppl:79–85.
39. Ramos-Goril JM, Pinto-Prades JL, Oppe M, Cabarrés JM, Semano-Agular P, Rivero-Arias O. Valuation and Modeling of EQ-5D-5L Health States Using a Hybrid Approach. *Med Care.* 2017;55:e51–8.
40. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Krol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34–42.
41. Beavens AS, Lounsbury JW, Richards JK, Huck SW, Skolits GJ, Esquivel SL. Practical considerations for using exploratory factor analysis in educational research. *Pract Assessment, Res Eval.* 2013;18:1–13.
42. Batista-Foguet JM, Coenders G, Nonso J. Análisis factorial confirmatorio. Su utilidad en la validación de cuestionarios relacionados con la salud. *Med. Clin. (Barc).* 2004;122:21–7.
43. Mulak SA, James LR, Van Alstine J, Bennett N, Lind S, Stilwell CD. Evaluation of goodness-of-fit indices for structural equation models. *Psychol Bull. American Psychological Association.* 1989;105:430.
44. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review. *J Educ Res Taylor & Francis.* 2006;99:323–38.
45. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16:297–334.
46. Cortina JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol.* 1993;78:98–104.
47. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics STOR.* 1977;33:362–74.
48. Katz LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care LWW.* 1989;27:5178–89.
49. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol.* 2003;56:395–407.
50. Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol Briefs.* 2004;57:1008–18.
51. Weir JP. The intraclass Correlation Coefficient and the SEM. *J Strength Cond Res.* 2005;19:231–40.
52. De Boer MR, De Vet HCW, Terwee CB, Moll AC, Wilker-Dieben HJM, Van Rens GHMB. Changes to the subscales of two vision-related quality of life questionnaires are proposed. *J Clin Epidemiol.* 2005;58:1260–8.
53. Harris K, Dawson J, Doll H, Field RE, Murray DW, Fitzpatrick R, et al. Can pain and function be distinguished in the Oxford Knee Score in a meaningful way? An exploratory and confirmatory factor analysis. *Qual Life Res Springer.* 2013;22:2561–8.
54. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med.* 1993;118:622–9.
55. Strand LI, Ljunggren AE, Bogen B, Ask T, Johnsen TB. The Short-Form McGill Pain Questionnaire as an outcome measure: Test-retest reliability and responsiveness to change. *Eur J Pain.* 2008;12:917–25.
56. Clement ND, MacDonald D, Simpson A. The minimal clinically important difference in the Oxford knee score and Short Form 12 score after total knee arthroplasty. *Knee Surgery, Sport, Traumatol Arthrosc. Spring.* 2014;22:1933–9.

57. Clement ND, Jenkins PJ, Macdonald D, Nie YX, Rutton JT, Breusch SJ, et al. Socioeconomic status affects the Oxford knee score and Short-Form 12 score following total knee replacement. *Bone Joint J.* 2013;95–B:52–8.
58. Luque R, Rizo B, Urdas A, Garcia-Creigo R, Moro E, Marco F, et al. Predictive factors for failure after total knee replacement revision. *Int Orthop.* 2014;38: 429–35.
59. Beswick AD, Wyde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients *BMJ Open.* 2012;2:e000435.

Submit your next manuscript to BioMed Central
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit



RESEARCH ARTICLE

Impact of lower limb osteoarthritis on health-related quality of life: A cross-sectional study to estimate the expressed loss of utility in the Spanish population

Jesús Martín-Fernández^{1,2,3*}, Roberto García-Maroto^{4,5}, Amaia Bilbao^{3,6,7}, Lidia García-Pérez^{3,8}, Blanca Gutiérrez-Teira⁹, Antonio Molina-Siguero¹⁰, Juan Carlos Arenaza^{3,11}, Vanesa Ramos-García⁸, Gemma Rodríguez-Martínez¹², Fco Javier Sánchez-Jiménez¹³, Gloria Ariza-Cardiel¹⁻³

1 Unidad Docente Multiprofesional de Atención Familiar y Comunitaria Oeste, Gerencia Asistencial de Atención Primaria, Servicio Madrileño de Salud, Madrid, Spain, **2** Facultad de Ciencias de la Salud, Universidad Rey Juan Carlos, Madrid, Spain, **3** Red de Investigación en Servicios Sanitarios y Enfermedades Crónicas (REDISSEC), Spain, **4** Servicio de Cirugía Ortopédica y Traumatología, Hospital Universitario Clínico San Carlos, Servicio Madrileño de Salud, Madrid, Spain, **5** Doctorando en el Programa de Investigación en Ciencias Médico Quirúrgicas, Facultad de Medicina, Universidad Complutense de Madrid, Madrid, Spain, **6** Osakidetza, Hospital Universitario Basurto, Unidad de Investigación, Bilbao, Spain, **7** Instituto de Investigación en Servicios de Salud Kronikgune, Barakaldo, Spain, **8** Fundación Canaria Instituto de Investigación Sanitaria de Canarias (FIISC), Las Palmas de Gran Canaria, Spain, **9** Centro de Salud El Soto, Gerencia Asistencial de Atención Primaria, Servicio Madrileño de Salud, Madrid, Spain, **10** Centro de Salud Presentación Sabio, Gerencia Asistencial de Atención Primaria, Servicio Madrileño de Salud, Madrid, Spain, **11** Osakidetza, Hospital Universitario Basurto, Servicio de Traumatología y Cirugía Ortopédica, Bilbao, Spain, **12** Centro de Salud Infante Don Luis, Gerencia Asistencial de Atención Primaria, Servicio Madrileño de Salud, Madrid, Spain, **13** Centro de Salud Gregorio Marañón, Gerencia Asistencial de Atención Primaria, Servicio Madrileño de Salud, Madrid, Spain

* jmfernandez@salud.madrid.org



OPEN ACCESS

Citation: Martín-Fernández J, García-Maroto R, Bilbao A, García-Pérez L, Gutiérrez-Teira B, Molina-Siguero A, et al. (2020) Impact of lower limb osteoarthritis on health-related quality of life: A cross-sectional study to estimate the expressed loss of utility in the Spanish population. PLoS ONE 15(1): e0228398. <https://doi.org/10.1371/journal.pone.0228398>

Editor: Yuanyuan Wang, Monash University, AUSTRALIA

Received: November 8, 2019

Accepted: January 14, 2020

Published: January 24, 2020

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0228398>

Copyright: © 2020 Martín-Fernández et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Objective

Osteoarthritis of the lower limb (OALL) worsens health-related quality of life (HRQL), but this impact has not been quantified with standardized measures. We intend to evaluate the impact of OALL on HRQL through measures based on individual preferences in comparison to the general population.

Methods

A cross-sectional study was designed. A total of 6234 subjects aged 50 years or older without OALL were selected from the Spanish general population (National Health Survey 2011–12). An opportunistic sample of patients aged 50 years or older diagnosed with hip ($n = 331$) or knee osteoarthritis ($n = 393$), using the American Rheumatism Association criteria, was recruited from six hospitals and 21 primary care centers in Vizcaya, Madrid and Tenerife between January and December 2015. HRQL was measured with the EQ-5D-5L, and the results were transformed into utility scores. Sociodemographic variables (age, sex, social group, cohabitation), number of chronic diseases, and body mass index were

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study has been funded by Instituto de Salud Carlos III through the Grants listed below. Co-funded by European Regional Development Fund, (ERDF) "A way of shaping Europe". JMF received the grant PI1300648, AB received the grant PI1300560, LGP received the grant PI1300518. The Foundation for Research and Biomedical Innovation of Primary Care of the Community of Madrid (FIBAP, by its initials in Spanish) has subsidized the translation and publication costs of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

considered. The clinical stage of OALL was collected using the Western Ontario and McMaster Universities Osteoarthritis Index and the Oxford hip score and Oxford knee score. Generalized linear models were constructed using the utility index as the dependent variable.

Results

HRQL expressed by OALL patients was significantly worse than this of the general population. After adjustment for sociodemographic and clinical characteristics, the mean utility loss was -0.347 (95% CI: -0.390, -0.303) for osteoarthritis of the hip and -0.295 (95% CI: -0.336, -0.255) for osteoarthritis of the knee. OALL patients who were treated at a hospital had an additional utility loss of -0.112 (95% CI: -0.158, -0.065).

Conclusion

OALL has a great impact on HRQL. People with OALL perceive a utility loss of approximately 0.3 points compared to the general population without osteoarthritis, which is very high in relation to the utility loss reported for other chronic diseases and for arthritis in general.

Introduction

Health is a fundamental component of well-being in developed societies. The definition of health proposed by the World Health Organization in 1948 surpassed the biological concept of illness and included aspects of physical, psychological and social well-being [1], making it necessary to incorporate new measures to evaluate health status. Health-related quality of life (HRQL) appears to be a necessary measure for assessing an individual's perceived well-being. Although there is no universally accepted definition of HRQL, it has been suggested that measurements of HRQL should incorporate the subject's perception of his or her health situation, the impact of the disease or its treatment on his or her well-being and how that impact affects his or her functionality [2].

The assessment of HRQL is fundamental in evaluating the impact of disease and health intervention outcomes on both the individual and society as whole. Measures based on patient individual preferences are essential tools in these evaluations [3]. They allow patients to describe the impact of poor health and to calculate the "utility" scores (or rates) associated with each description of health status, making it possible to attribute value to such states through the study of preferences.

The most widely used tool for the measurement of HRQL in Spain is the EQ-5D [4], a questionnaire based on individual preferences that, through an algorithm, allows the attribution of a "utility" score to each described health state. This questionnaire is the most widely used tool for analyzing the cost-effectiveness of health care technologies [5]. Some European health technology assessment organizations, such as the National Institute for Health and Care Excellence (NICE) in the United Kingdom, have specifically stated that the EQ-5D is the preferred measure of HRQL in adults; thus, utilizing it ensures comparability among studies [6,7]. In Spain, the scores or utilities derived from the latest version of this tool, the EQ-5D-5L, have been proposed to provide information on economic evaluations of technologies [8].

The assessment of perceived health status in patients with chronic diseases is essential for measuring the impact and burden of disease, as in the case of patients with osteoarthritis of the lower limb (OALL) affecting the hip and knee. OALL represents a widely recognized public health burden. It is the eleventh most common cause of disability worldwide, and the disability-adjusted life years (DALYs) lost to it increased by 34.8% from 2005 to 2015 [2]. In Spain, a prevalence of 45.0% for knee osteoarthritis and 24.1% for hip osteoarthritis has been reported in people over 65 years of age [10]. Annual costs of €1,500 were estimated in 2007 for patients with osteoarthritis of the knee or hip in Spain, of which 86% were direct costs [11]. Other studies estimate costs of approximately €5,000 per year for patients in Europe and €12,000 per year in the US (2013 euros) [12]. Currently, the health care costs for generalized osteoarthritis account for between 0.25% and 0.50% of the Spanish gross domestic product (GDP) [13]. Furthermore, it is expected that the prevalence of OALL will continue to increase in the future, due largely to the aging of the population [14] and the increased prevalence of obesity [15]. As a result, this condition will likely continue to present substantial challenges for health planning in the coming years. Reliable measurements of health outcomes in OALL are increasingly important for health decision makers, health professionals and patients [16]. An essential element for advancing knowledge of the impact of OALL is the availability of measures that can be used to accurately and thoroughly measure the individual burden posed by this disease [17].

In Spain, HRQL has been assessed in patients with OALL who consulted trauma specialists [18] and in older people with OALL in the community setting [19], and the results suggest that this disease is associated with a marked decrease in perceived well-being; however, to date, no study has determined what differences exist between this group and the rest of the community to isolate the effects of the disease from those produced by concomitant sociodemographic or clinical conditions.

This study aims to evaluate the impact of OALL on HRQL through measures based on individual preferences in comparison to the general population that does not have osteoarthritis.

Materials and methods

Design

A cross-sectional study was designed. The included population was collected from a random sample of the Spanish general population and from a sample of patients with OALL. The general population data were obtained by selecting the population aged 50 years or older from the 2011–2012 National Health Survey (NHS), of which the methodological bases are publicly available [20]. Of the 21,007 available surveys, 6,234 responses of subjects 50 years or older who stated that they did not have a medical diagnosis of osteoarthritis, arthritis, or rheumatism were chosen. To study patients with OALL, information was obtained through an opportunistic sampling of patients diagnosed with osteoarthritis of the hip or knee according to the American Rheumatism Association criteria [21,22] who consulted traumatology and rheumatology specialists at six hospitals and 21 primary care centers in Vizcaya, Madrid and Tenerife between January and December 2015. Patients with malignant or other organic diseases or psychiatric disorders that hindered participation and those who could neither read nor understand Spanish were excluded. Of the total subjects included ($n = 758$), only those 50 years of age or older were selected ($n = 724$).

Data regarding people with OALL were collected directly from the patients, and data regarding the general population were extracted from the NHS.

This sample size allowed the construction of explanatory models that were appropriate for achieving the proposed objective [23].

Dependent variable

To measure the perceived HRQL in both samples, the EQ-5D-5L was used [4]. In its most current version, the EQ-5D-5L [24] consists of two parts: a 0-to-100 scale to assess HRQL, the Visual Analog Scale (VAS) and a questionnaire comprising 5 questions or domains (mobility, self-care, carrying out usual activities, pain/discomfort, and anxiety/depression) with 5 response levels (ranging from 1, no problems, to 5, impossibility or severe problems). From these five questions, 3,125 health states are obtained, and the score associated with each state is the utility index. Initially, it was considered that the utility index should range between 0, a state equivalent to death, and 1, which represented perfect health. However, individual preference studies of different health states have identified states that are less preferable than death for the general population, which means this index can yield negative values (up to -0.421) using the algorithms published for Spain [8]. The reliability, validity and sensitivity of the EQ-5D-5L have been studied in patients with OALL in Spain [25].

Independent variables

The age and sex of each subject and his or her social group, based on a six-category classification related to occupation (with group I being the highest and group VI the lowest) [26], were collected as sociodemographic variables. Information on cohabitation was included, and those who lived alone were differentiated from those living with someone (in any form of cohabitation).

Perceived health status was assessed using a Likert scale of five categories (very good, good, fair, poor and very poor) and the number of chronic diseases diagnosed. The patients' body mass index (BMI) and whether they had undergone a primary care or hospital consultation were also studied.

Only for the OALL group was it possible to determine whether there was unilateral or bilateral involvement and whether osteoarthritis had been diagnosed in the other large lower limb joint (hip or knee). The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [27] and the Oxford knee score (OKS) [28] and Oxford hip score (OHS) [29] questionnaires were used to characterize the clinical stage of OALL. The WOMAC is a multidimensional scale comprising 24 items measuring the domains of pain (five items), stiffness (two items) and physical function (17 items) in patients with OALL [27]. We used the version with five response levels, scored from 0 to 4, representing different degrees of intensity (none, mild, moderate, severe or extreme) for each item. Those scores are summed and standardized to yield a score from 0 to 100 (from better to worse capacity). The higher the score, the worse the patient's status is. This questionnaire has been adapted and validated in Spain [30]. The OKS and OHS scales measure the severity of symptoms in patients with osteoarthritis of the knee and hip, respectively. Each scale consists of 12 questions, and the scores, which range from 0 to 48 points, classify the clinical situation of osteoarthritis patients in 4 groups: excellent (> 41 points), good (between 34 and 41 points), moderate (between 27 and 34 points) and poor (<27 points) [31]. The Spanish versions of the OKS and OHS questionnaires have also been validated [32,33].

Analysis

We present the descriptive statistics of the explanatory and dependent variables with their measures of central tendency and dispersion. The qualitative variables were compared using chi-squared tests, and the quantitative variables were compared using Student's t test or, if necessary, its nonparametric equivalent.

To address the main objective, generalized linear models (GLMs) were constructed, with the utility index attributable to the subject's perceived health status as the dependent variable. To select the best model, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were examined. The Gaussian family and "identity" link function were selected as the more appropriate ones using these criteria. Standard errors (SEs) were calculated using robust methods to prevent possible heteroscedasticity [3,34].

Model 1 included variables related to the reported HRQL and the presence of OALL, model 2 added the affected joint, and model 3 also included whether the patient had been evaluated in the hospital setting. Improvements in model fit were compared by calculating Akaike weights, which express the probability that a new model is better than the set of tested models [35]. The improvement in the BIC was also studied according to the interpretations proposed by Kass and Raftery [36]. The variable "perceived health status" was not included as an explanatory variable due to the risk of overfitting when measuring the same construct as the dependent variable. Neither bilaterality nor having another joint group affected by OALL (hip or knee) improved the explanatory power of the final model (model 3). Patients with any missing data were not included in the models. The results section explains the number of subjects included in each model.

Results

The results are presented for 6958 subjects aged 50 or older: 6234 from the general population without osteoarthritis and 724 (10.41% of the total) with a diagnosis of OALL. Of the patients diagnosed with OALL, 393 (54.28%) had a diagnosis of osteoarthritis of the knee, of whom 53 (13.49%) had a previous diagnosis of osteoarthritis of the hip. Of the 724 patients with OALL, 331 (45.72%) had a diagnosis of osteoarthritis of the hip, of whom 115 (34.74%) had a previous diagnosis of osteoarthritis of the knee. A total of 286 patients with OALL were recruited from primary care clinics (39.50%), and 438 (60.50%) were recruited from specialized care practices: 369 (50.97%) from traumatology and 69 (9.53%) from rheumatology practices.

Table 1 shows the characteristics of the total sample studied. Among the patients with OALL, women, people older than 65 years, and more disadvantaged social groups predominated. There was also a higher prevalence of overweight and obesity and of cohabitation with some type of partner.

The perceived health status and utility expressed by patients with OALL were significantly worse than those expressed by the general population in terms of the transformation of the EQ-5D-5L into utilities and the VAS (Table 1) and of each domain of the questionnaire (Table 2). The differences in HRQL were most pronounced for the domains mobility, carrying out usual activities, and pain/discomfort. The range of utilities for the expressed health states ranged from -0.416 (state 55555) to 1 (state 11111) for both the general population and the patients with OALL. In the case of patients with osteoarthritis of the knee, the worst health state reported was 44555, which has a utility value of -0.297. The responses to the VAS of the EQ-5D-5L ranged from 0 to 100 for both the general population and the patients with OALL (both hip and knee).

Table 3 shows the level of severity for patients with osteoarthritis of the knee or hip as measured with the WOMAC and the OHS/OKS [30,32,33]. Those with osteoarthritis of the knee more frequently had bilateral involvement, but those with osteoarthritis of the hip more commonly showed concomitant involvement of the other large lower limb joint. There were no significant differences between the patients with knee and hip osteoarthritis in the utilities or the level of severity measured with specific instruments. Only the VAS of the EQ-5D-5L

Table 1. Characteristics of the studied sample.

	General population N = 6234	Population with OALL N = 724	p
Age (%)			
50–54 years	21.82	4.70	<0.001
55–59 years	18.53	8.15	
60–64 years	15.77	13.12	
65–69 years	13.97	16.57	
70–74 years	9.62	17.54	
75–79 years	8.87	20.99	
80–84 years	6.26	13.40	
>85 years	5.17	5.52	
Age, mean (SD)	64.44 (10.81)	70.93 (9.13)	<0.001
Sex (%)			
Female	47.06	62.71	<0.001
Social group (%)			
Group I	11.64	7.12	<0.001
Group II	7.93	4.56	
Group III	19.00	16.79	
Group IV	15.22	20.62	
Group V	32.02	29.38	
Group VI	14.20	21.53	
Cohabitation (%)			
With partner	61.93	66.99	<0.001
Chronic diseases (%)			
None	17.85	0.00	<0.001
One	22.44	55.11	
Two	20.82	30.39	
Three or more	38.88	14.50	
Chronic diseases, mean (SD)	2.34 (2.02)	1.66 (0.91)	<0.001
BMI (%)			
Underweight	1.03	0.14	<0.001
Normal	31.18	19.06	
Overweight	40.02	43.23	
Obese	27.77	37.57	
BMI, mean (SD)	26.69 (4.09)	29.02 (4.72)	<0.001
Perceived health status (%)			
Very good	13.60	2.23	<0.001
Good	55.25	22.67	
Fair	23.24	48.26	
Poor	6.42	21.56	
Very poor	1.49	5.29	
EQ-5D-5L, utilities, mean (SD)	0.924 (0.160)	0.532 (0.287)	<0.001
EQ-5D-5L, VAS, mean (SD)	74.71 (17.71)	56.31 (21.70)	<0.001

BMI: Body Mass Index; SD: Standard Deviation; VAS: Visual Analog Scale.

<https://doi.org/10.1371/journal.pone.0228398.t001>

showed a trend of patients with arthritis of the hip presenting worse perceived HRQL than those diagnosed with osteoarthritis of the knee, although the difference was not significant.

Table 4 shows the results of the explanatory models for the expressed utilities. Model 3 emerged as the best model because its Akaike weights, compared to those of the set of models

Table 2. Distribution of the responses to the different domains of the EQ-5-5L for the general population and the population with lower limb osteoarthritis (OALL).

EQ-5D-5L domain	Percentage of responses per level					P
	1	2	3	4	5	
Mobility						
General population n = 6234	84.62	7.92	4.23	2.29	0.93	<0.001
OALL n = 723	10.93	20.06	43.02	23.93	2.07	
Self-care						
General population n = 6234	93.55	2.98	1.44	0.90	1.12	<0.001
OALL n = 722	28.39	27.70	31.58	11.36	0.97	
Usual activities						
General population n = 6234	89.09	5.21	2.68	1.57	1.44	<0.001
OALL n = 722	16.48	25.21	36.70	16.62	4.99	
Pain/discomfort						
General population n = 6231	76.46	14.35	6.55	2.33	0.32	<0.001
OALL n = 723	6.22	21.30	37.07	30.57	4.84	
Anxiety/depression						
General population n = 6226	85.21	9.14	3.82	1.48	0.35	<0.001
OALL n = 716	44.13	23.18	18.30	11.31	3.07	

<https://doi.org/10.1371/journal.pone.0228398.t002>

and model 2, had a value of 1; that is, there is statistical certainty that it is the best model. In terms of improvement in the BIC, the evidence against the goodness of fit of models 1 and 2 was very strong (with BIC differences well above 10 in both cases). The three models, especially model 3, considerably reduced the error variance, as shown by the value of McFadden's adjusted R^2 .

Table 3. Characteristics of patients with OALL included in the study.

	Patients with hip osteoarthritis n = 331	Patients with knee osteoarthritis n = 393	p
Bilateral involvement (%)	28.40	42.24	<0.001
Involvement of other large lower limb joint (hip or knee) (%)	34.74	13.49	<0.001
State according to Oxford score (%)			
Excellent	4.56	3.32	0.348
Good	11.85	8.42%	
Moderate	18.24	19.64	
Poor	65.35	68.32	
EQ-5D-5L utilities, mean (SD)	0.517 (0.303)	0.543 (0.272)	0.110
EQ-5D-5L VAS, mean (SD)	54.92 (21.78)	57.47 (21.59)	0.059
Oxford score, mean (SD)	22.72 (10.58)	21.97 (9.94)	0.163
WOMAC score, mean (SD)			
WOMAC score, pain	45.82 (22.63)	47.09 (20.55)	0.216
WOMAC score, limitation	52.65 (22.81)	51.14 (20.99)	0.178
WOMAC score, stiffness	48.33 (25.97)	46.81 (25.29)	0.213
WOMAC score, total	50.85 (22.02)	49.95 (20.10)	0.282

SD: Standard Deviation; VAS: Visual Analog Scale. WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

<https://doi.org/10.1371/journal.pone.0228398.t003>

Table 4. Explanatory models of the differences in utility derived from the EQ-5D-5L.

Variable	Model 1 Coef (CI 95%)	Model 2 Coef (CI 95%)	Model 3 Coef (CI 95%)
Level of care			
Hospital care vs other			-0.112 (-0.158, -0.065)
Lower limb OA			
Hip vs none		-0.406 (-0.444, -0.369)	-0.347 (-0.390, -0.303)
Knee vs none		-0.365 (-0.396, -0.334)	-0.295 (-0.336, -0.255)
Lower limb OA			
Yes vs No	-0.384 (-0.409, -0.360)	-	-
Age			
55–59 vs 50–54 years	-0.004 (-0.014, 0.006)	-0.004 (-0.014, 0.006)	-0.004 (-0.014, 0.006)
60–64 vs 50–54 years	0.007 (-0.003, 0.017)	0.007 (-0.003, 0.017)	0.007 (-0.003, 0.018)
65–69 vs 50–54 years	0.013 (0.002, 0.024)	0.013 (0.002, 0.024)	0.013 (0.002, 0.024)
70–74 vs 50–54 years	0.006 (-0.008, 0.020)	0.005 (-0.009, 0.019)	0.005 (-0.009, 0.019)
75–89 vs 50–54 years	0.004 (-0.012, 0.019)	0.003 (-0.012, 0.018)	0.003 (-0.012, 0.018)
80–84 vs 50–54 years	-0.062 (-0.084, -0.040)	-0.062 (-0.084, -0.040)	-0.064 (-0.086, -0.042)
≥85 vs 50–54 years	-0.148 (-0.182, -0.114)	-0.148 (-0.182, -0.114)	-0.150 (-0.184, -0.117)
Sex			
Male vs female	0.021 (0.014, 0.029)	0.022 (0.014, 0.030)	0.022 (0.014, 0.030)
Social class			
Group II vs group I	-0.008 (-0.022, 0.006)	-0.009 (-0.022, 0.005)	-0.010 (-0.024, 0.004)
Group III vs group I	-0.021 (-0.034, -0.008)	-0.021 (-0.034, -0.008)	-0.021 (-0.034, -0.008)
Group IV vs group I	-0.017 (-0.029, -0.004)	-0.017 (-0.030, -0.004)	-0.017 (-0.030, -0.004)
Group V vs group I	-0.023 (-0.034, -0.012)	-0.023 (-0.035, -0.012)	-0.023 (-0.035, -0.012)
Group VI vs group I	-0.024 (-0.037, -0.010)	-0.024 (-0.038, -0.010)	-0.024 (-0.038, -0.010)
Cohabitation			
With partner vs alone	0.010 (0.002, 0.018)	0.010 (0.002, 0.019)	0.010 (0.002, 0.019)
Chronic diseases			
One vs none	-0.012 (-0.020, -0.005)	-0.012 (-0.020, -0.004)	-0.010 (-0.018, -0.003)
Two vs none	-0.028 (-0.037, -0.019)	-0.028 (-0.037, -0.019)	-0.030 (-0.039, -0.021)
Three or more vs none	-0.094 (-0.103, -0.085)	-0.094 (-0.102, -0.085)	-0.093 (-0.102, -0.084)
BMI			
Underweight vs nonobese/obese weight	-0.048 (-0.102, 0.006)	-0.047 (-0.101, 0.007)	-0.047 (-0.100, 0.006)
Overweight vs nonobese	0.005 (-0.004, 0.013)	0.004 (-0.004, 0.013)	0.004 (-0.005, 0.012)
Obese vs nonobese	-0.028 (-0.039, -0.017)	-0.028 (-0.039, -0.018)	-0.029 (-0.040, -0.019)
Characteristics of the model			
Family: Gaussian	N = 6542	N = 6542	N = 6542
Link function: Identity	AIC = -5522.47	AIC = -5529.77	AIC = -5593.02
Fami	BIC = -5373.18	BIC = -5373.69	BIC = -5430.15
	McFadden's $R^2_a = 0.549$	McFadden's $R^2_a = 0.553$	McFadden's $R^2_a = 0.562$

CI: Confidence Interval; BMI: Body Mass Index.

<https://doi.org/10.1371/journal.pone.0228398.t004>

Model 3 shows how the presence of hip osteoarthritis decreases utility by an average of 0.347 points for people who are similar in terms of age, sex, social group, state of cohabitation and burden of chronic illness. For patients with knee osteoarthritis, the decrease in this value is 0.295 points on average. For any type of osteoarthritis, being treated in the hospital environment is associated with an average decrease in utility of 0.112 points, adjusted for the characteristics mentioned above.

Factors associated with a lower level of utility, in addition to OALL and being treated in a hospital environment, included being older (over 80 years), belonging to a more disadvantaged social group, being female, being obese, living alone, and presenting more chronic conditions.

Discussion

This is the first study in Spain to quantify the impact of OALL on HRQL through preference measures adjusted for variables that are known to be associated with changes in perceived HRQL, such as age, number of chronic diseases, sex and other social conditions, such as social group or potential loneliness. In addition, an algorithm to estimate the social rating or individual preferences for health states was used to measure this impact according to the latest version of the EQ-5D-5L [8], a tool with advantages over previous versions and whose Spanish version has shown excellent psychometric properties for patients with OALL [25]. OALL is associated with a substantial decrease in the utility attributed to the state of health when this association is adjusted for other, potentially confounding variables. Living with OALL implies a utility loss of approximately 0.30 on average compared to the general population without osteoarthritis, which exceeds the thresholds of clinical relevance obtained in Spain (0.07 points in the EQ-5D-5L utility index for subjects with nonsurgically treated OALL at the group level) [25]. On the other hand, as could be inferred, patients treated at the specialized level had an even worse perception of their HRQL, and their attributed utilities averaged 0.11 points lower.

Degenerative joint involvement was reported by the general population in the United Kingdom as one of the chronic conditions that most affects quality of life, behind only pain and anxiety/depression [37]. Similar data have been reported for the Canadian population [38]. The utility loss attributed to osteoarthritis in these studies has generally been approximately 0.10 points when adjusted for other variables that could affect perceived HRQL, a loss approximately three times higher than that produced by other chronic conditions, such as diabetes or asthma [38]. In Spain, when the EQ-5D-3L was used to evaluate HRQL, the study of osteoarthritis as a whole identified it as a chronic condition that decreases utility to the same extent as cardiac problems or diabetes mellitus, with a decrease in utility of up to 0.10 points compared to the general population over 65 years of age [39]. The literature has reported that OALL has significant impacts on utility, similar to those found in this study. In an Italian population, an average loss of utility of approximately 0.28 points was observed in patients with OALL compared to the general population [40]. In this case, utility was evaluated with the 3-level version of the EQ-5D, and adjustments were made only for age and sex. According to these authors, the impact of OALL on utilities was similar to that produced by osteoporosis with vertebral fracture or ankylosing spondylitis and was much higher than that produced by other chronic diseases that affect the musculoskeletal system, such as Sjogren's syndrome or systemic sclerosis [40]. These figures far surpass the sensitivity thresholds of the EQ-5D-5L obtained in Spain. The established minimal clinically important difference (MCID) at the group level in Spanish patients with nonsurgically treated OALL was 0.07 [25], and in other countries, this threshold was 0.09 for subjects with arthritis [38].

It has also been possible to study the association between the setting from which the patient is recruited and his or her health situation. The average utility score for patients treated in the hospital setting was 0.11 points lower. In Spain, HRQL has been assessed in OALL patients who consulted trauma specialists [18] and in those in the community [19]; however, the results were not comparable due to the use of different measurement tools. This is the first time that evidence has been presented regarding the different degrees of disease involvement among patients seen at different levels of care, and the results are consistent with the organization of the Spanish national health system, in which the first level of care serves as a gateway for services for different chronic diseases and for osteoarthritis in particular [41].

It is possible that the great impact of OALL on HRQL occurs because the disease affects several of the domains incorporated into the EQ-5D tool, such as pain and loss of function [42]. However, in this case, the data presented show the impairment of a psychological component captured in the anxiety/depression domain of the EQ-5D-5L. This impact of OALL on mental health has previously been described in the literature [43]. In addition, the mental health domain as a component of perceived HRQL has been shown to be associated with joint replacement in OALL [44], so it should be evaluated with special attention.

On the other hand, the results of this study do not allow us to say that patients with osteoarthritis of the hip have, in general, a worse perception of their state of health than those with osteoarthritis of the knee, as has been suggested in some studies [18]. Although this tendency could be suspected from the descriptive analysis of the data (the VAS of the EQ-5D-5L), it is not observed when the models are adjusted, and the confidence intervals of the coefficients of hip/knee osteoarthritis overlap considerably.

The remaining sociodemographic and clinical characteristics, which were adjusted to determine the association between OALL and the utilities of health states, behaved in ways that were previously known. Older people, those from more disadvantaged socioeconomic groups, those who live alone, women, obese people and those with more chronic conditions tended to have a worse perceived health status, as described in previous studies in Spain [45,46]. It has previously been shown that the association between osteoarthritis and reported HRQL is confounded by other variables, such as sociodemographic factors, chronic disease or obesity [19,37]; therefore, the results presented here can be considered a realistic approach to the study of the association between OALL and HRQL.

Limitations

The design of this study has some limitations. Cross-sectional studies are problematic when establishing causal associations, although the effect of the main confounding factors collected in the literature has been assessed. Additionally, the sample of patients with OALL cannot be understood as representative of the population with this condition but was collected using opportunity criteria. We can affirm that sampling was carried out in diverse geographical sites and that the profile of the patients with OALL (predominantly female, older, and with a lower socioeconomic status and a higher prevalence of obesity) coincides with that reported in other Spanish [47] and European [19] studies. The fact that the participants may have had more severe disease than patients in the general population due to overrepresentation in the hospital setting was taken into account in the analysis. The general population from whom data were collected comprised people who did not report any type of osteoarthritis or arthritis diagnosis. The approach to the existence of chronic conditions could be considered a limitation of the study, as the conditions, or lack thereof, were reported by the subjects. However, even though there are problems with the general population data, the NHS constitutes the most representative and highest-quality information available at the time.

The EQ-5D-5L is a generic instrument designed to measure dimensions of health relevant to all health states, including healthy individuals, and not patients' perception of aspects of health specifically affected by OALL. However, it has the advantage of allowing us to attribute patients' preferences to the described health status in patients with OALL and to compare them between different populations [48].

Implications

The importance of the results presented lies in the interest that health policy decision makers may have in reproducible and comparable quantifications of the impact of OALL on HRQL.

These results quantify the burden of the disease from the perspective of the patient. In addition, there are interventions, such as joint replacement surgeries, that involve a considerable investment of resources but can provide very valuable results from the perspective of the patient and society as a whole. It has frequently been noted that to assess the outcome of such interventions, it is necessary to use patient-reported outcome measures, such as quality of life, that help determine how to prioritize actions [49].

Conclusion

OALL is a chronic disease that has a great impact on the HRQL of patients. Patients with OALL perceive a very significant loss of utility, approximately 0.3 points compared to the general population without osteoarthritis. This impact differs depending on the place where the patient has been treated. Patients treated in the hospital setting in a health system in which primary care functions as a gateway for health care services report an additional utility loss of 0.1 point. This utility loss attributable to OALL is very high in relation to what has been reported for other chronic diseases and for arthritis in general, exceeding the threshold of the so-called MCID by up to three times.

Supporting information

S1 Data. Data for Sharing.

(XLS)

S1 File. STROBE-checklist_cross-sectional.

(DOC)

Acknowledgments

We are grateful to colleagues in the participating hospitals and primary care centers for their support and to all patients for their collaboration.

Author Contributions

Conceptualization: Jesús Martín-Fernández, Roberto García -Maroto, Amaia Bilbao, Lidia García-Pérez, Gloria Ariza-Cardiel.

Data curation: Jesús Martín-Fernández, Amaia Bilbao, Lidia García-Pérez, Blanca Gutiérrez-Teira, Antonio Molina-Siguero, Juan Carlos Arenaza, Fco Javier Sánchez-Jiménez, Gloria Ariza-Cardiel.

Formal analysis: Jesús Martín-Fernández, Amaia Bilbao.

Funding acquisition: Jesús Martín-Fernández, Amaia Bilbao, Lidia García-Pérez, Gloria Ariza-Cardiel.

Investigation: Jesús Martín-Fernández, Roberto García -Maroto, Amaia Bilbao, Lidia García-Pérez, Blanca Gutiérrez-Teira, Antonio Molina-Siguero, Juan Carlos Arenaza, Vanesa Ramos-García, Gemma Rodríguez-Martínez, Fco Javier Sánchez-Jiménez, Gloria Ariza-Cardiel.

Methodology: Jesús Martín-Fernández, Amaia Bilbao, Lidia García-Pérez, Vanesa Ramos-García, Gloria Ariza-Cardiel.

Project administration: Jesús Martín-Fernández, Amaia Bilbao, Lidia García-Pérez, Gemma Rodríguez-Martínez, Gloria Ariza-Cardiel.

Visualization: Jesús Martín-Fernández, Roberto García-Maroto.

Writing – original draft: Jesús Martín-Fernández, Amaia Bilbao.

Writing – review & editing: Jesús Martín-Fernández, Roberto García-Maroto, Amaia Bilbao, Lidia García-Pérez, Blanca Gutiérrez-Teira, Antonio Molina-Siguero, Juan Carlos Arenaza, Vanesa Ramos-García, Gemma Rodríguez-Martínez, Fco Javier Sánchez-Jiménez, Gloria Ariza-Cardiel.

References

1. World Health Organization. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19–22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into the force on 7 April 1948.
2. Urzúa A. Calidad de vida relacionada con la salud: Elementos conceptuales. *Rev Med Chile*. 2010; 138: 358–365.
3. Walloo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, et al. Mapping to estimate health-state utility from non-preference-based outcome measures: An ISPOR good practices for outcomes research task force report. *Value Health*. 2017; 20: 18–27. <https://doi.org/10.1016/j.jval.2016.11.006> PMID: 28212961
4. Group EuroQol. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy*. 1990; 16: 199–208. [https://doi.org/10.1016/0168-8510\(90\)90421-9](https://doi.org/10.1016/0168-8510(90)90421-9) PMID: 10109801
5. Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabases JM, Serrano-Aguilar P, Rivero-Arias O. Valuation and modeling of EQ-5D-5L health states using a hybrid approach. *Med Care*. 2017; 55: e51–e58.6. <https://doi.org/10.1097/MLR.000000000000283> PMID: 25521503
6. NICE National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. London; 2013.
7. Dakin H. Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. *Health Qual Life Outcomes*. 2013; 11: 151. <https://doi.org/10.1186/1477-7525-11-151> PMID: 24010873
8. Ramos-Goñi JM, Craig BM, Oppe M, Ramallo-Farina Y, Pinto-Prades JL, Luo N, et al. Handling data quality issues to estimate the Spanish EQ-5D-5L value set using a hybrid interval regression approach. *Value Health*. 2018; 21: 596–604. <https://doi.org/10.1016/j.jval.2017.10.023> PMID: 29753358
9. Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*. 2016; 388: 1545–1602. [https://doi.org/10.1016/S0140-6736\(16\)31678-6](https://doi.org/10.1016/S0140-6736(16)31678-6) PMID: 27733282
10. van der Pas S, Castel MV, Cooper C, Denkiner M, Dennison EM, Edwards MH, et al. European project on osteoarthritis: Design of a six-cohort study on the personal and societal burden of osteoarthritis in an older European population. *BMC Musculoskelet Disord*. 2013; 14: 138. <https://doi.org/10.1186/1471-2474-14-138> PMID: 23597054
11. Loza E, Lopez-Gomez JM, Abasolo L, Maese J, Carmona L, Batlle-Gualda E. Economic burden of knee and hip osteoarthritis in Spain. *Arthritis Rheum*. 2009; 61: 158–165. <https://doi.org/10.1002/art.24214> PMID: 19177521
12. Salmon JH, Rat AC, Sellam J, Michel M, Eschard JP, Guillemin F, et al. Economic impact of lower-limb osteoarthritis worldwide: A systematic review of cost-of-illness studies. *Osteoarthritis Cartilage*. 2016; 24: 1500–1508. <https://doi.org/10.1016/j.joca.2016.03.012> PMID: 27034093
13. Puig-Junoy J, Zamora AR. Socio-economic costs of osteoarthritis: A systematic review of cost-of-illness studies. *Semin Arthritis Rheum*. 2015; 44: 531–541. <https://doi.org/10.1016/j.semarthrit.2014.10.012> PMID: 25511476
14. Prieto-Alhambra D, Judge A, Javald MK, Cooper C, Diez-Perez A, Arden NK. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: Influences of age, gender and osteoarthritis affecting other joints. *Ann Rheum Dis*. 2014; 73: 1659–1664. <https://doi.org/10.1136/annrheumdis-2013-203355> PMID: 23744977
15. Reyes C, Leyland KM, Peat G, Cooper C, Arden NK, Prieto-Alhambra D. Association between overweight and obesity and risk of clinically diagnosed knee, hip, and hand osteoarthritis: A population-based cohort study. *Arthritis Rheumatol*. 2016; 68: 1869–1875. <https://doi.org/10.1002/art.39707> PMID: 27059260

16. Ackerman IN, Busija L, Tacey MA, Bohensky MA, Ademi Z, Brand CA, et al. Performance of the assessment of quality of life measure in people with hip and knee joint disease and implications for research and clinical use. *Arthritis Care Res (Hoboken)*. 2014; 66: 481–488.
17. Busija L, Osborne RH, Roberts C, Buchbinder R. Systematic review showed measures of individual burden of osteoarthritis poorly capture the patient experience. *J Clin Epidemiol*. 2013; 66: 826–837. <https://doi.org/10.1016/j.jclinepi.2013.03.011> PMID: 23810023
18. Bernad-Pineda M, de las 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: 283–289. <https://doi.org/10.1016/j.recot.2014.04.005> PMID: 25022212
19. van Schoor NM, Zambon S, Castell MV, Cooper C, Denkiner M, Dennison EM, et al. Impact of clinical osteoarthritis of the hip, knee and hand on self-rated health in six European countries: The European project on OsteoArthritis. *Qual Life Res*. 2016; 25: 1423–1432. <https://doi.org/10.1007/s11136-015-1171-8> PMID: 26547441
20. Instituto Nacional de Estadística. Ministerio de Sanidad Servicios Sociales e Igualdad. Metodología de la Encuesta Nacional de Salud 2011–2012. [Internet]. 2013 pp. 1–2. Available from: https://www.msssi.gob.es/estadEstudios/estadisticas/encuestaNacional/encuestaNac2011/MetodologiaENSE2011_12.pdf
21. Altman RD. Criteria for classification of clinical osteoarthritis. *J Rheumatol Suppl*. 1991; 27: 10–12. PMID: 2027107
22. 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: 1039–1049. <https://doi.org/10.1002/art.1780290816> PMID: 3741515
23. Green SB. How many subjects does it take to do a regression analysis. *Multivariate Behav Res*. 1991; 26: 499–510. https://doi.org/10.1207/s15327906mbr2603_7 PMID: 26776715
24. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011; 20: 1727–1736. <https://doi.org/10.1007/s11136-011-9903-x> PMID: 21479777
25. Bilbao A, García-Pérez L, Arenaza JC, García I, Ariza-Cardel G, Trujillo-Martín E, et al. Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: Reliability, validity and responsiveness. *Qual Life Res*. 2018; 27: 2897–2908. <https://doi.org/10.1007/s11136-018-1929-x> PMID: 29978346
26. Domingo SA, Marcos AJ. Proposal of an indicator of "social class" based on the occupation. *Gac Sanit*. 1989; 3: 320–326. [https://doi.org/10.1016/s0213-9111\(89\)70948-1](https://doi.org/10.1016/s0213-9111(89)70948-1) PMID: 2498222
27. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol*. 1988; 15: 1833–1840. PMID: 3068365
28. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br*. 1998; 80: 63–69. <https://doi.org/10.1302/0301-620X.80B1.7859> PMID: 9460955
29. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg Br*. 1996; 78: 185–190. PMID: 8666621
30. Escobar A, Quintana JM, Bilbao A, Azkarate J, Guenaga JI. Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. Western Ontario and McMaster universities osteoarthritis index. *Clin Rheumatol*. 2002; 21: 466–471. <https://doi.org/10.1007/s100670200117> PMID: 12447629
31. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, et al. The use of the Oxford hip and knee scores. *J Bone Joint Surg Br*. 2007; 89: 1010–1014. <https://doi.org/10.1302/0301-620X.89B8.19424> PMID: 17785736
32. Martín-Fernández J, García-Maroto R, Sánchez-Jiménez FJ, Bau-González A, Valencia-García H, Gutiérrez-Teira B, et al. Validation of the Spanish version of the Oxford knee score and assessment of its utility to characterize quality of life of patients suffering from knee osteoarthritis: A multicentric study. *Health Qual Life Outcomes*. 2017; 15: 186. <https://doi.org/10.1186/s12955-017-0761-2> PMID: 28962613
33. Martín-Fernández J, Gray-Laymon P, Molina-Siguero A, Martínez-Martín J, García-Maroto R, García-Sánchez I, et al. Cross-cultural adaptation and validation of the Spanish version of the Oxford hip score in patients with hip osteoarthritis. *BMC Musculoskelet Disord*. 2017; 18: 205. <https://doi.org/10.1186/s12891-017-1568-3> PMID: 28532445

34. Long JS, Ervin LH. Using heteroscedasticity consistent standard errors in the linear regression model. *Am Stat*. 2000; 54: 217–224.
35. Fabozzi F, Focardi S, Rachev S, Arshanapalli B. Appendix E: Model selection criterion: AIC and BIC. In: Fabozzi F, Focardi S, Rachev S, Arshanapalli B, editors. *The basics of financial econometrics: Tools, concepts, and asset management applications*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2014. pp. 399–430.
36. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc*. 1995; 90: 773–795.
37. Wu M, Brazier JE, Kearns B, Relton C, Smith C, Cooper CL. Examining the impact of 11 long-standing health conditions on health-related quality of life using the EQ-5D in a general population sample. *Eur J Health Econ*. 2015; 16: 141–151. <https://doi.org/10.1007/s10198-013-0559-z> PMID: 24408476
38. Tsiplova K, Pullenayegum E, Cooke T, Xie F. EQ-5D-derived health utilities and minimally important differences for chronic health conditions: 2011 commonwealth fund survey of sicker adults in Canada. *Qual Life Res*. 2016; 25: 3009–3016. <https://doi.org/10.1007/s11136-016-1336-0> PMID: 27307010
39. Oliva-Moreno J, Lopez-Bastida J, Worbes-Cerezo M, Serrano-Aguilar P. Health related quality of life of Canary Island citizens. *BMC Public Health*. 2010; 10: 675. <https://doi.org/10.1186/1471-2458-10-675> PMID: 21054832
40. Salaffi F, Di Carlo M, Carotti M, Farah S, Ciapetti A, Gutierrez M. The impact of different rheumatic diseases on health-related quality of life: A comparison with a selected sample of healthy individuals using SF-36 questionnaire, EQ-5D and SF-6D utility values. *Acta Biomed*. 2019; 89: 541–557. <https://doi.org/10.23750/abm.v89i4.7298> PMID: 30657123
41. Ministerio de Sanidad, Servicios Sociales e Igualdad. Estrategia en enfermedades reumáticas y musculoesqueléticas del Sistema Nacional de Salud. Madrid, Spain: Ministerio de Sanidad, Servicios Sociales e Igualdad; 2013.
42. Carlesso LC, Hawker GA, Waugh EJ, Davis AM. Disease-specific pain and function predict future pain impact in hip and knee osteoarthritis. *Clin Rheumatol*. 2016; 35: 2999–3005. <https://doi.org/10.1007/s10067-016-3401-z> PMID: 27601231
43. Veronese N, Stubbs B, Solmi M, Smith TO, Noale M, Cooper C, et al. Association between lower limb osteoarthritis and incidence of depressive symptoms: Data from the osteoarthritis initiative. *Age Ageing*. 2017; 46: 470–476. <https://doi.org/10.1093/ageing/afw216> PMID: 27932358
44. Serra-Sutton V, Allepuz A, Martínez O, Espallargues M. Factores relacionados con la calidad de vida al año de la artroplastia total de cadera y rodilla: Estudio multicéntrico en Cataluña. *Rev Esp Cir Ortop Traumatol*. 2013; 57: 254–262. <https://doi.org/10.1016/j.recot.2013.04.002> PMID: 23885650
45. Martín-Fernández J, Ariza-Cardiel G, Polentinos-Castro E, Sanz-Cuesta T, Sarria-Santamera A, Del Cura-Gonzalez I. Explaining differences in perceived health-related quality of life: A study within the Spanish population. *Gac Sanit*. 2018; 32: 447–453. <https://doi.org/10.1016/j.gaceta.2017.05.016> PMID: 28958573
46. Arrospide A, Machon M, Ramos-Giori JM, Ibarrondo O, Mar J. Inequalities in health-related quality of life according to age, gender, educational level, social class, body mass index and chronic diseases using the Spanish value set for Euroqol 5D-5L questionnaire. *Health Qual Life Outcomes*. 2019; 17: 69. <https://doi.org/10.1186/s12955-019-1134-9> PMID: 30999899
47. Carou AC, Fernandez SP, Diaz SP, Santos FJD. Clinical profile, level of affection and therapeutic management of patients with osteoarthritis in primary care: The Spanish multicenter study EVALUA. *Rheumatol Clin*. 2015; 11: 353–360. <https://doi.org/10.1016/j.reuma.2014.12.005> PMID: 25648903
48. García-Pérez L, Ramos-García V, Serrano-Aguilar P, Pais-Brito JL, Aciego de Mendoza M, Martín-Fernández J, et al. EQ-5D-5L utilities per health states in Spanish population with knee or hip osteoarthritis. *Health Qual Life Outcomes*. 2019; 17: 164. <https://doi.org/10.1186/s12955-019-1230-x> PMID: 31666071
49. Torre JA, Puigdomenech E, Valderas JM, Evans JP, Martín V, Molina AJ, et al. Availability of specific tools to assess patient reported outcomes in hip arthroplasty in Spain. Identifying the best candidates to incorporate in an arthroplasty register. A systematic review and standardized assessment. *PLoS One*. 2019; 14: e0214746. <https://doi.org/10.1371/journal.pone.0214746> PMID: 30934024