


RESEARCH ARTICLE

Open Access



# Minimal clinically important decline in physical function over one year: EPOSA study

Paola Siviero<sup>1\*</sup> , Federica Limongi<sup>1</sup>, Antonella Gesmundo<sup>2</sup>, Sabina Zamboni<sup>2</sup>, Cyrus Cooper<sup>3</sup>, Elaine M. Dennison<sup>3</sup>, Mark H. Edwards<sup>3,4</sup>, Erik J. Timmermans<sup>5</sup>, Suzan van der Pas<sup>5</sup>, Laura A. Schaap<sup>5</sup>, Natasja M. van Schoor<sup>5</sup>, Michael D. Denkiner<sup>6</sup>, Florian Herbolzheimer<sup>7</sup>, Richard Peter<sup>7</sup>, Maria Victoria Castell<sup>8</sup>, Ángel Otero<sup>8</sup>, Rocio Queipo<sup>8</sup>, Nancy L. Pedersen<sup>9</sup>, Dorly J. H. Deeg<sup>5</sup>, Stefania Maggi<sup>1</sup> and for the EPOSA Research Group

## Abstract

**Background:** The Australian/Canadian hand Osteoarthritis Index (AUSCAN) and the Western Ontario and McMaster Universities knee and hip Osteoarthritis Index (WOMAC) are the most commonly used clinical tools to manage and monitor osteoarthritis (OA). Few studies have as yet reported longitudinal changes in the AUSCAN index regarding the hand. While there are published data regarding WOMAC assessments of the hip and the knee, the two sites have always evaluated separately. The current study therefore sought to determine the minimal clinically important difference (MCID) in decline in the AUSCAN hand and WOMAC hip/knee physical function scores over 1 year using anchor-based and distribution-based methods.

**Methods:** The study analysed data collected by the European Project on Osteoarthritis, a prospective observational study investigating six adult cohorts with and without OA by evaluating changes in the AUSCAN and WOMAC physical function scores at baseline and 12–18 months later. Pain and stiffness scores, the performance-based grip strength and walking speed and health-related quality of life measures were used as the study's anchors. Receiver operating characteristic curves and distribution-based methods were used to estimate the MCID in the AUSCAN and WOMAC physical function scores; only the data of those participants who possessed paired (baseline and follow up-measures) AUSCAN and WOMAC scores were included in the analysis.

**Results:** Out of the 1866 participants who were evaluated, 1842 had paired AUSCAN scores and 1845 had paired WOMAC scores. The changes in the AUSCAN physical function score correlated significantly with those in the AUSCAN pain score ( $r = 0.31$ ). Anchor- and distribution-based approaches converged identifying 4 as the MCID for decline in the AUSCAN hand physical function. Changes in the WOMAC hip/knee physical function score were significantly correlated with changes in both the WOMAC pain score ( $r = 0.47$ ) and the WOMAC stiffness score ( $r = 0.35$ ). The different approaches converged identifying two as the MCID for decline in the WOMAC hip/knee physical function.

**Conclusions:** The most reliable MCID estimates of decline over 1 year in the AUSCAN hand and WOMAC hip/knee physical function scores were 4 and 2 points, respectively.

**Keywords:** Osteoarthritis, MCID, Functional decline, Hand, Hip/knee, AUSCAN, WOMAC

\* Correspondence: [paola.siviero@in.cnr.it](mailto:paola.siviero@in.cnr.it)

<sup>1</sup>National Research Council, Neuroscience Institute - Aging Branch, Via Giustiniani 2, 35128 Padova, Italy

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



## Background

The Australian/Canadian Hand Osteoarthritis Index (AUSCAN) [1] and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [2, 3] scales are self-report instruments measuring pain, stiffness and physical function linked to osteoarthritis (OA), and have been used by the European Project on Osteoarthritis (EPOSA) to assess personal and societal variables affected by OA, such as quality of life (QoL), social participation, and health care use in several ageing European cohorts. The individuals enrolled in the project were receiving treatment for severe OA, had undiagnosed or untreated OA or did not have OA at all [4].

The Minimum Clinically Important Difference (MCID) is defined as the smallest change in a score that a patient perceives as beneficial or detrimental [5]. There are different types of MCID, depending on whether there has been an improvement or a worsening in the variable being measured and on the external standard being employed [6].

Until now, to our knowledge, the MCID in the AUSCAN hand and WOMAC hip/knee physical function scales has received scarce attention. Specifically few studies report longitudinal changes in the AUSCAN [7]. While some studies have investigated the WOMAC scales [7–9], the two sites of hip and knee have always been evaluated separately [10–15]. Moreover, the MCID has almost always been considered from an improvement perspective, as the majority of studies have aimed to examine the efficacy of pharmacological interventions [11, 13], rehabilitation programs [8, 9], and/or of surgical treatments [10, 12, 14, 15].

The aim of the current study was therefore to estimate the MCID in the AUSCAN and WOMAC physical function subscales using distribution-based and anchor-based methods for longitudinal changes. We postulated that the AUSCAN and the WOMAC physical function scores would worsen [16, 17] with time (i.e., there would be a rise in both) and that the changes in the AUSCAN and WOMAC physical function scores would correlate significantly with changes in other well-established OA health variables or performance-based measures [18–25].

## Methods

### Participants

The current study analysed data collected by the European Project on Osteoarthritis (EPOSA), a population-based study involving cohorts living in Germany, Italy, the Netherlands, Spain, Sweden, and the UK, that recruited 2942 adults between the ages of 65–85. All the participants gave written informed consent; the study design and methodology are outlined in detail elsewhere [4]. The study design was granted approval by the appropriate local ethics committees (Germany:

Universität Ulm Ethikkommission [312/08]. Italy: Comitato Etico Provinciale Treviso [XLIV-RSA/AULSS7]. The Netherlands: Medisch Ethische Toetsingscommissie Vrije Universiteit Amsterdam [2002/141]. Spain: Comité Ético de Investigación Clínica del Hospital Universitario La Paz Madrid [PI-1080]. Sweden: Till forskningsetikommittén vid Karolinska Institutet Stockholm [00–132]. UK: Hertfordshire Research Ethics Committee [10/H0311/59]).

The project aimed to evaluate the participants once at baseline (between November 2010 and November 2011) and a second time 12–18 months later. During the assessment the participants underwent a clinical examination and were interviewed at home or in a health care centre by trained physicians and nurses using a standardized questionnaire.

### Measures

*Physical function, pain and stiffness of hand OA* were assessed at baseline and 12–18 months later using the three subscales of the AUSCAN (15 items grouped into 3 scales: pain (5 items), stiffness (1 item), and physical function (9 items)) that utilized a 5-point Likert scale (responses ranged from none to extreme; 0 = none, 1 = mild, 2 = moderate, 3 = severe, and 4 = extreme) [1].

*Physical function, pain and stiffness of hip and/or knee OA* were measured at baseline and 12–18 months later using the three subscales of the WOMAC (24 items grouped into 3 scales: pain (5 items), stiffness (2 items), and physical function (17 items)) that utilized a 5-point Likert scale. Hip/knee pain and stiffness were defined as the maximum value of two joints [2, 3].

All the AUSCAN and WOMAC subscales were normalized from 0 to 100; higher scores indicate worse health status [1–3].

*The Grip strength* of both hands at baseline and 12–18 months later was measured twice, using a strain gauge dynamometer and the result (the highest values of the right and left hands were summed and divided by 2) is expressed in kilograms [26].

*Walking speed* was measured by time, registered in seconds, for a 3-m course marked out on the floor with no obstructions for an additional 2 ft at both ends.

*Anxiety and depression* were evaluated at baseline and 12–18 months later using the Hospital Anxiety and Depression Scales (HADS), a 14-item self-assessment instrument that measures anxiety and/or depression separately [27]. Scores that are 8 or higher (scores range between 0 and 21) on each/either of the subscales indicate altered states.

*Health-related QoL* was measured at baseline and 12–18 months later using the 5-level EQ-5D, consisting of a descriptive system comprising five dimensions (mobility, self care, usual activities, pain/discomfort, anxiety/

depression) and the EQ VAS, a vertical visual analogue scale [28]. The scores of the EQ-5D were converted into a single index, based on general population surveys, using the time trade-off (TTO) valuations from the general population of the UK; scores between  $-0.594$  and  $1$  indicate good or satisfying health. Scores of the EQ-VAS range between  $0$  and  $100$ , with higher scores indicating better health.

*Clinical diagnosis of OA* was formulated on the basis of the participant's medical history and a physical examination (only at baseline), utilising algorithms in accordance with the clinical criteria developed by the American College of Rheumatology (ACR) [29] and the recommendations of the European League Against Rheumatism [30].

Clinical hand OA (classified as present vs absent) was diagnosed using specific AUSCAN sections [1]: the cut-off in the algorithm for hand pain was  $\geq 3$  and it was  $\geq 1$  for stiffness. At least 2 of the following criteria were required for a diagnosis of hand OA: a) hard tissue enlargement of two or more joints, b) hard tissue enlargement of two or more distal inter-phalangeal joints, c) deformity of at least one hand joint. Swelling of the metacarpophalangeal joints criteria was a variable that was assessed only in the English and German participants.

Clinical hip/knee OA, defined as the presence of OA in at least one or both of these joints, was diagnosed on the basis of the outcome of specific WOMAC sections. Pain in the hip/knee on at least one side [2, 3] was evaluated during the physical examination using a cut-off  $\geq 3$ . For the participants, to be diagnosed with knee OA, at least two of the following were necessary: a) a morning stiffness score from mild to extreme; b) crepitus with active motion on at least one side at the physical examination; c) bone tenderness on at least one side at the physical examination; d) bone enlargement at the physical examination on at least one side; e) no palpable warmth of synovium at the physical examination in either knees. All of the following were needed for a positive hip OA assessment: a) pain in the hip on at least one side associated with restricted hip internal rotation at a physical examination; b) morning stiffness of the hip lasting  $< 60$  min, evaluated using the stiffness section of the WOMAC with a score from mild to extreme.

### Statistical analysis

Data analyses and graphical presentations were carried out using SAS software (SAS System, SAS Institute Inc., Cary, NC), version 9.4. Data were analysed using a set of weights calculated per sex and per 5-year age class with respect to the 2010 Standard European Population [4].

The changes over times (in the 12–18 months between the baseline evaluation and the follow-up one) were evaluated as continuous variables using the non-parametric

signed rank test. Spearman's correlation was used to compare the changes in the AUSCAN and WOMAC physical function scales and the changes in the other variables; the Cronbach  $\alpha$  coefficient was used to measure the scales' reliability (internal consistency) (values of  $\alpha \geq 0.7$  reflect a good reliability) [31].

Only the data of the participants whose assessments were considered complete, that is they had completed both the baseline and follow-up assessments, were included in the statistical analysis. The MCID was calculated by measuring the changes from basal to follow-up measurements scores. Since the MCID for subject-reported outcome measures may vary in different populations and depending on the context, as recommended by Revicki et al. [32], we used multiple approaches to estimate the MCID in the AUSCAN and WOMAC physical function scores to triangulate on a single value or on a small range of values.

For **anchor-based estimation** of MCID we used the receiver operating characteristic (ROC) curve on the change score in the anchor. The variables assessed as possible anchors for the AUSCAN hand OA physical function score were: the AUSCAN for hand OA Pain, the AUSCAN for hand OA Stiffness, the Grip strength, the HADS anxiety, the HADS depression, the EQ-5D-5 L, and the EQ VAS. The variables evaluated as possible anchors for the WOMAC for hip/knee OA physical function score were: the WOMAC for hip/knee OA Pain, the WOMAC for hip/knee OA Stiffness, the Walking-test time, the HADS anxiety, the HADS depression, the EQ-5D-5 L, and the EQ VAS.

An anchor should be chosen because of a significant correlation between the change in the physical function score and the change in the anchor and a correlation coefficient  $\geq |0.30|$  [31].

A ROC curve was constructed for those participants showing stable or worsened anchor scores; the area under the curve (AUC) summarizes the instrument's ability to distinguish between individuals who have or do not have a minimal clinically important difference in functionality. The criteria used to calculate the probability of an optimal cut-off were: the Youden index (J) [33], the Euclidean distance (D), and the equality sensitivity and specificity (S). The percentage of participants exceeding the MCID were estimated for each cut-off value.

The following were considered for the **distribution based-methods**:

1. A standardized response mean (SRM) [34], representing very small (SRM2), moderate (SRM5) and large changes (SRM8) [35].
2. The standard error of measurement (SEM) [36] of the changes, considering a 63% confidence interval

(CI) (SEM63), a 90% CI (SEM90), and a 95% CI (SEM95).

- The Edwards-Nunnally (EN) method [37], at the 90% CI (EN90), and at the 95% CI (EN95).

Anchor-based and distribution-based-methods were used to determine the MCID, and on that basis the participants were divided into two categories: worse/no worse functionality at the second assessment point. Triangulation was used to examine multiple values from different approaches to converge on a single value, with Cohen’s k (range from -1 to 1, with one indicating a perfect agreement).

### Results

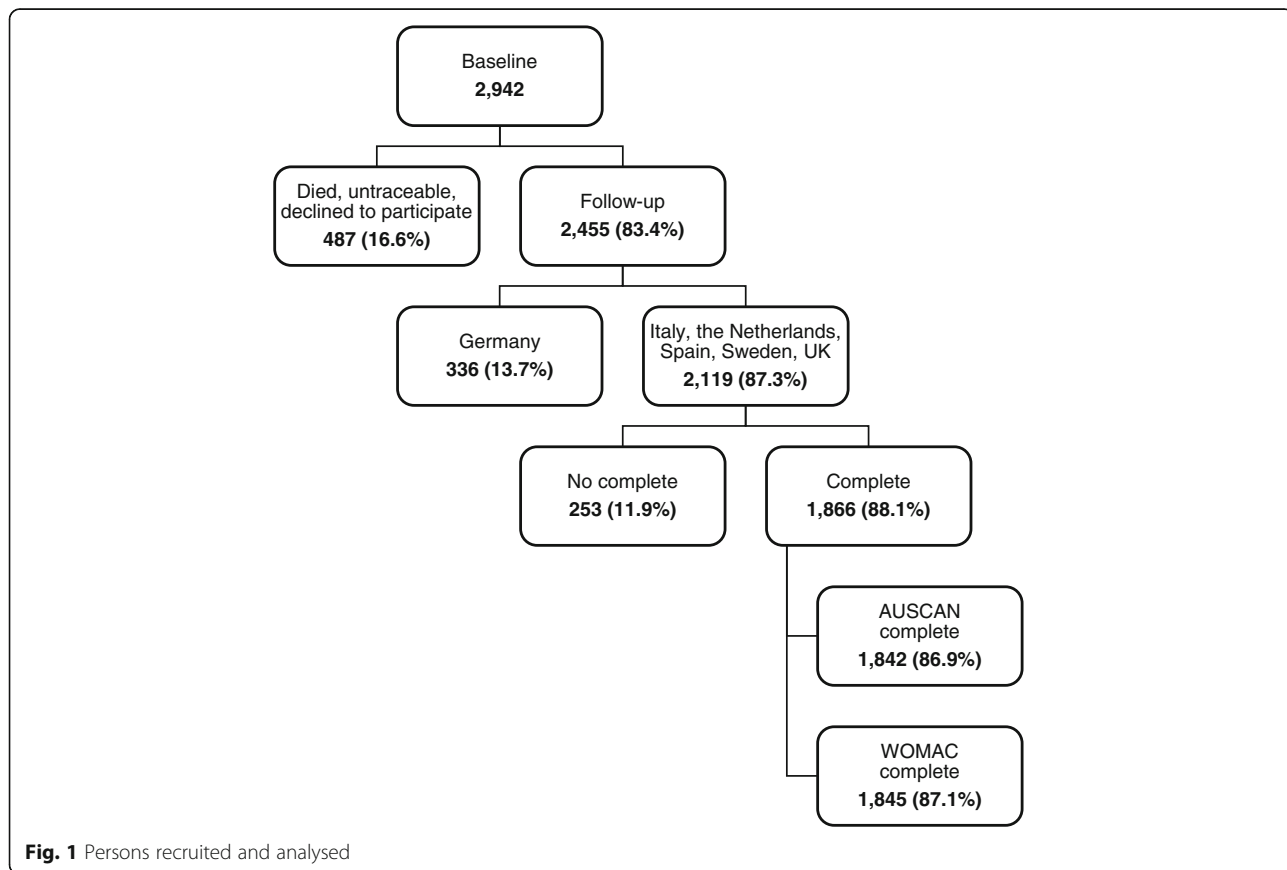
Out of the original 2942 participants who completed the baseline evaluation (Fig. 1), 2455 (83%) agreed to undergo a follow-up evaluation 12–18 months later. It was not possible to re-evaluate 487 participants (16.6% of the baseline sample) because they had died, were untraceable, or declined to participate. The non-completers were significantly older, more likely to be female, less educated, and predominantly Italian with respect to the completers.

Since information from the German group was incomplete, data from that country ( $n = 336$ , 14% of 2455), were not analysed. The data that was analysed and upon which our results are therefore based represents 1866 participants who completed both of the study’s evaluations. One thousand, eight hundred forty-two of these had paired AUSCAN measurements and 1845 had paired WOMAC measurements.

Table 1 (weighted data) shows that approximately 17% of the participants had clinical hand OA and more than 22% had clinical hip/knee OA. The median changes in the physical function scores detected using the AUSCAN hand and WOMAC hip/knee subscales were significant, as were the median changes in the WOMAC stiffness score. There were no significant changes in the AUSCAN pain and stiffness subscales and in the WOMAC pain subscale over time.

All other changes in grip strength, walking-test time, the HADS scales, and the EQ-5D-5 L were significant over time.

Table 2 shows the correlation coefficient in the change in the AUSCAN and the WOMAC physical function scores and in the change in the other measures that were considered as possible anchors. For the hand, only the changes in the AUSCAN pain scores were significantly



**Fig. 1** Persons recruited and analysed

**Table 1** Baseline characteristics and change in physical function and in the anchors after 12–18 months

	Baseline		Change: 12–18 months follow-up - baseline			<i>P</i>	
Age, n, mean (SD), median (IQR), y	1886	73.8 ± 5.0	73 (70.77)				
Female Sex, n (%)	1886	977 (51.8)					
High education, n (%)	1886	1120 (58.8)					
Country, n (%)	1886						
Italy		312 (16.8)					
The Netherlands		407 (22.0)					
Spain		413 (22.4)					
Sweden		392 (23.0)					
UK		362 (15.8)					
Clinical hand OA, n (%)	1886	1842 (16.7)					
Clinical hip/knee OA, n (%)	1886	1845 (22.6)					
AUSCAN for hand OA <sup>a</sup> , n, mean ± SD, median (IQR)							
Physical function	1842	9.1 ± 15.4	0 (0,11)	1842	1.1 ± 11.3	0 (0,4)	<b>&lt;.0001</b>
Pain <sup>c</sup>	1842	7.6 ± 15.2	0 (0,5)	1842	1.7 ± 14.4	0 (0,0)	0.1949
Stiffness	1842	10.0 ± 18.7	0 (0,25)	1842	1.0 ± 17.2	0 (0,0)	0.1358
WOMAC for hip/knee OA <sup>a</sup> , n, mean ± SD, median (IQR)							
Physical function	1845	8.5 ± 13.2	2 (0,12)	1845	1.5 ± 11.3	0 (-1,4)	<b>0.0123</b>
Pain	1845	10.0 ± 14.3	5 (0,15)	1845	1.4 ± 13.2	0 (-5,5)	0.2579
Stiffness	1845	12.9 ± 18.6	0 (0,25)	1845	0.3 ± 17.9	0 (0,12)	<b>0.0390</b>
Grip strength <sup>b</sup> , n, mean ± SD, median (IQR), kg	1862	27.7 ± 10.2	26 (20,35)	1826	-1.1 ± 5.8	-1 (-3.5,1.5)	<b>0.0004</b>
Walking-test time <sup>c</sup> , n, mean ± SD, median (IQR), sec	1854	3.3 ± 1.6	3 (2.3,3.8)	1830	0.1 ± 1.6	0 (-0.5,0.6)	<b>&lt;.0001</b>
HADS anxiety <sup>d</sup> , n, mean ± SD, median (IQR)	1886	4.5 ± 3.5	4 (2,7)	1872	-0.6 ± 2.6	-1 (-2,1)	<b>&lt;.0001</b>
HADS depression <sup>d</sup> , n, mean ± SD, median (IQR)	1886	3.5 ± 3.0	3 (1,5)	1875	-0.3 ± 2.4	0 (-1,1)	<b>&lt;.0001</b>
EQ-5D-5 L <sup>e</sup> , n, mean ± SD, median (IQR)	1876	0.8 ± 0.2	0.8 (0.7,1.0)	1855	0.0 ± 0.2	0 (0.0,0.1)	<b>0.0059</b>
EQ VAS <sup>f</sup> , n, mean ± SD, median (IQR)	1882	76.5 ± 16.2	80 (70,90)	1875	-0.5 ± 15.6	0 (-10,5)	<b>0.0114</b>

Weighted data. Numbers of subjects, age, and sex were unweighted data. Except where indicated otherwise, values are the percent of subjects EPOSA European Project on OsteoArthritis, SD Standard deviation, IQR interquartile range, AUSCAN AUStralian/CANadian Osteoarthritis Hand Index, WOMAC Western Ontario and McMaster Universities Osteoarthritis Index, OA osteoarthritis, HADS Hospital Anxiety and Depression Scales, EQ-5D-5 L Health status using five dimensions, EQ VAS Health status using the visual analogue scale

<sup>a</sup> Possible scores range from 0 to 100, with 100 indicating worse health status. <sup>b</sup> Lower values indicate worse performance. <sup>c</sup> Higher values indicate worse performance. <sup>d</sup> Possible scores range from 0 to 21, a score ≥ 8 indicates anxiety and/or depression. <sup>e</sup> Possible scores range from -0.594 to 1, lower values indicating worse health status. <sup>f</sup> Possible scores range from 0 to 100, with 0 indicating worse health status

The *P* are bold where they are less than or equal to the significance level cut-off of 0.05

correlated with a coefficient greater than |0.3| ( $r = 0.31$ ) with the changes in the AUSCAN physical function scores. For the hip/knee, both the changes in the WOMAC pain scores and the changes in the WOMAC stiffness scores were correlated with the changes in the WOMAC physical function score with respectively a correlation coefficient of 0.47 and 0.35 (greater than |0.3|).

The AUSCAN and WOMAC physical function scores showed a very good reliability of coefficient (Cronbach's  $\alpha$  of 0.92 and 0.94 respectively).

#### AUSCAN physical function estimates of MCID

Using hand pain as an external anchor, the estimates of the MCID in the AUSCAN hand physical function were consistent and equal to one. The only divergent criteria was the Youden index according to which the estimated

MCID for the hand was four. Using distribution-based methods, the estimate for significant worsening in the AUSCAN physical function score ranged from 1 to 8.

Based on these cut-offs, the participants were divided into worse vs not worse in functionality 12–18 months after baseline (Fig. 2).

When the percentage of values obtained with distribution-based MCID methods were compared with those produced by anchor-based methods, the two sets agreed most strongly according to Cohen's  $k$ . The first set ( $k$  values at approximately 0.97) formed by the ROC D Euclidean distance, the ROC S for equal sensitivity and specificity, and the SRM2, identified approximately 34–35% of the participants with clinically significant physical function decline at the 12–18 month follow-up evaluation. The second set ( $k$  values ranging from 0.95

**Table 2** Correlation among change in physical function with change in the anchors

Change Anchors	AUSCAN for hand OA Physical function (n = 1842)		WOMAC for hip/knee OA Physical function (n = 1845)	
	r	P	r	P
AUSCAN for hand OA Pain	0.30742	<.0001		
AUSCAN for hand OA Stiffness	0.21075	<.0001		
WOMAC for hip/knee OA Pain			0.46557	<.0001
WOMAC for hip/knee OA Stiffness			0.34937	<.0001
Grip strength	-0.10,451	<.0001		
Walking-test time			0.06917	0.0034
HADS anxiety	0.02249	0.3363	0.11065	<.0001
HADS depression	0.04346	0.0630	0.12391	<.0001
EQ-5D-5 L	-0.11449	<.0001	-0.15293	<.0001
EQ VAS	-0.07177	0.0021	-0.10105	<.0001

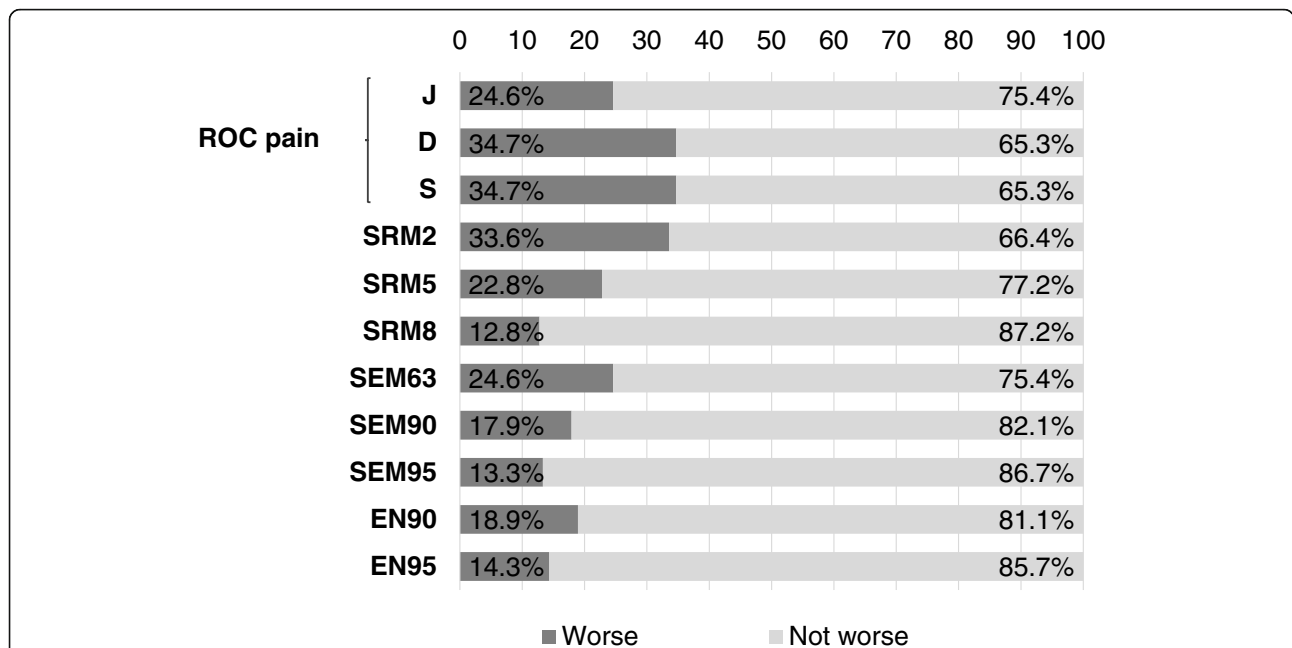
AUSCAN AUStralian/CANadian Osteoarthritis Hand Index, WOMAC Western Ontario and McMaster Universities Osteoarthritis Index, OA osteoarthritis, HADS Hospital Anxiety and Depression Scales, EQ-5D-5 L health status using five dimensions, EQ VAS health status using the visual analogue scale, r Spearman's correlation coefficient

to 1) which was formed by the ROC J Youden index, the SRM5, and the SEM63, uncovered that 24% of the participants had a clinically significant decline. In view of the concordance and the recommendation to privilege the anchor-based methods [32], we compared the MCID based on the ROC J/SEM63 and the one based on the ROC D/S. Out of the 639 worse participants identified by the ROC D/S criterion, 453 were the same ones identified by ROC J/SEM63. The MCID based on the ROC

J/SEM63, which estimated a change of 4 points, was found to be the most reliable criterion to analyse the loss of hand functionality at 12–18 months.

**WOMAC physical function estimates of MCID**

Using hip/knee pain and stiffness scales as external anchors, the estimates of the MCID of the WOMAC hip/knee physical function were consistent, although the magnitude of the correlations can only be considered



**Fig. 2** Estimates of MCID percentages for the AUSCAN hand physical function. AUSCAN: Australian/Canadian Hand Osteoarthritis Index; MCID: Minimum Clinically Important Difference; ROC pain: receiver operating characteristic using the AUSCAN pain score as the anchor; J: point that maximizes the Youden index; D: point that minimizes the Euclidean distance; S: point that minimizes the equality sensitivity, specificity. SRM: standardized response mean; SRM2: SRM with Cohen's threshold 0.20; SRM5: SRM with Cohen's threshold 0.50; SRM8: SRM with Cohen's threshold 0.80; CI: Confidence Interval; SEM: Standard Error Measurement; SEM63: SEM with 63% CI; SEM90: SEM with 90% CI; SEM95: SEM with 95% CI; EN: Edwards-Nunnally index; EN90: EN with 90% CI; EN95: EN with 95% CI

moderate. The ROC analysis of the anchor responses to the WOMAC pain and stiffness scales estimated that the MCID was almost always one. Once again, the divergent criteria was the Youden index with stiffness as the anchor that estimated a two point MCID for hip/knee physical function.

Using distribution-based methods, the estimate for significant worsening in the WOMAC physical function score ranges from 1 to 9. Figure 3 shows the percentage of participants who had worse hip/knee functionality 12–18 months after baseline according to the different methods utilized. The MCIDs for hip/knee physical function decline that showed the highest degree of agreement were: those based on the SRM2, those that used the WOMAC pain score as the anchor minimized the Euclidean distance and the equality sensitivity and specificity as well as those that maximized the Youden index ( $k = 0.94$ ). The SRM2 also agreed with those that, using the WOMAC stiffness score as the anchor, minimized the Euclidean distance and the equality sensitivity and specificity ( $k = 0.94$ ), or maximized the Youden index ( $k = 1$ ). These methods identified 30 and 33% of participants with clinically significant hip/knee physical function decline 12–18 months after baseline respectively. Finally, there was a strong agreement ( $k = 0.89$ ) between the SEM63 and the Youden index with the

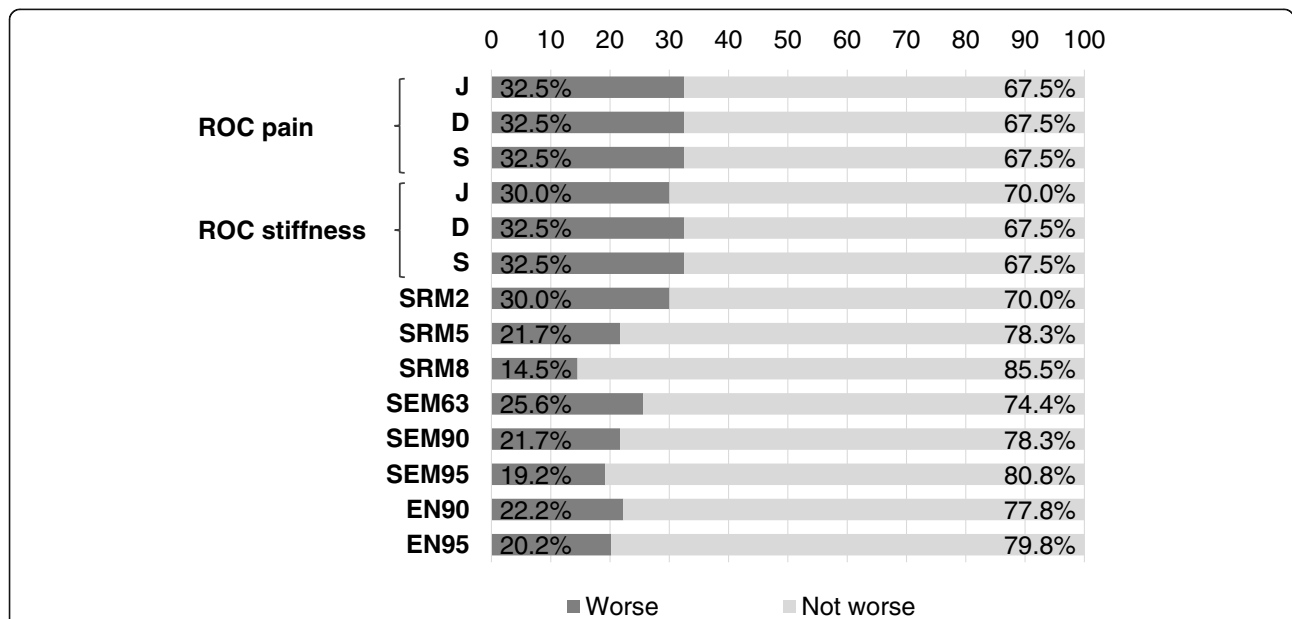
stiffness score used as an anchor; they respectively detected 26 and 30% of the participants.

As the highest degree of agreement was found between the Youden index (using ROC with stiffness as the anchor) and the SRM2, the MCID based on these criteria seemed to be the most suitable one to analyse the loss of hip/knee functionality at 12–18 months. Both criteria were consistent in identifying two as the best discriminating WOMAC physical function change cut-off.

**Discussion**

While it is true that the AUSCAN and WOMAC scales are the most commonly used clinical tools to manage and monitor OA patients, to our knowledge the MCID for decline picked up by these measures has never been evaluated. Only the study of Angst et al. [9], which focused on patients with OA of the lower extremities after a rehabilitation intervention, reported a MCID showing worsening in the WOMAC hip/knee physical function of approximately 1.33, based on a scale from 0 to 10. The study’s initial premise that hand and hip/knee physical function would deteriorate significantly over a year’s time was confirmed by our data showing higher AUSCAN and WOMAC physical function scores.

Although the relevance of the MCID approach remains controversial and despite the fact that physical



**Fig. 3** Estimates of MCID percentages for the WOMAC hip/knee physical function. WOMAC, Western Ontario and McMaster Universities; MCID: Minimum Clinically Important Difference; ROC pain: receiver operating characteristic using the WOMAC pain score as the anchor; ROC stiffness: receiver operating characteristic using the WOMAC stiffness score as the anchor; J: point that maximizes the Youden index; D: point that minimizes the Euclidean distance; S: point that minimizes the equality sensitivity, specificity. SRM: standardized response mean; SRM2: SRM with Cohen’s threshold 0.20; SRM5: SRM with Cohen’s threshold 0.50; SRM8: SRM with Cohen’s threshold 0.80; CI: Confidence Interval; SEM: Standard Error Measurement; SEM63: SEM with 63% CI; SEM90: SEM with 90% CI; SEM95: SEM with 95% CI; EN: Edwards-Nunnally index; EN90: EN with 90% CI; EN95: EN with 95% CI

function values can depend on the population being examined, the context, the time and methods used [25], it remains an important assessment instrument. The magnitude of the MCID was inferior in the participants studied here using the WOMAC instrument with respect to other studies [7–15]. As those studies focused on patients before and after interventions, the differences in the magnitude of the MCID might be connected to patient expectations regarding surgical interventions, as compared to non-surgical interventions [38]. Other factors that might explain the differing MCID values could be: the severity of the participant's baseline health status, the length of the period being examined, the accuracy of the measurement instruments, and the direction in the change in the MCID (i.e. towards improvement or worsening).

Other studies have demonstrated that changes in the AUSCAN and WOMAC physical function scores correlate significantly with changes in other generic, adapted, or performance-based measures used to gauge pain and function in the hand and hip/knee OA [18–25]. Our study did not, however, find any correlations between the changes in the AUSCAN and WOMAC physical function and the changes in other more generic measures such as the Hospital Anxiety and Depression Scale and the European QoL Surveys. The anchors with the strongest correlations were the pain-specific questionnaires (AUSCAN and WOMAC pain subscale), presumably because they are basically measures of pain during physical activities rather than unspecific pain measures [39].

The estimates of the MCID in both the AUSCAN hand physical function and the WOMAC hip/knee physical function according to the ROC analysis using different anchor responses and criteria, were consistent. The only divergent criteria was the Youden index that overestimated the MCIDs.

But as explained above, besides an anchor-based method, we also used a distribution-based approach to estimate the MCID, given that the two are complementary [40]. Distribution-based approaches, which are based on the statistical characteristics of the samples studied and reliable measures, generated a wide range of different estimates of the MCID for the AUSCAN hand and WOMAC hip/knee physical function that were greater than the anchor-based estimates. Both methods converged to a common result.

While distribution-based estimates are able to furnish supportive information when the change is significant, they do not provide a direct measure of minimum clinically important difference. That is why precedence was shown to the anchor based estimates. Moreover, since the MCIDs estimated using distribution-based methods were greater than the mean change reported 12–18 months after baseline, it is possible that the data from

the distribution-based methods provide information about clinical significance but might overestimate the true MCID.

This study has several limitations. First, the changes in outcome measures could hypothetically be associated with baseline levels. Second, there is the possibility that the participants selected did not experience much or any change over the 12–18 month study period. Third, there may be even important differences in the populations studied and in the cut-off values of the MCID physical function decline. Indeed, estimates can differ depending on the instrument, domain, country, and condition, at least for condition-specific measures, and further research is required before the estimates presented here can be generalized to other instruments [7].

The study's most important strength was undeniably its large population base: the participants were randomly selected from older community-dwelling European populations. Not only persons with OA but also large numbers of healthy individuals not affected with OA were analysed. The methodology used was the same in all of the countries, and OA was diagnosed in accordance with standardized international guidelines [4]. Our study was based on valid standardized globalized measures (the WOMAC and AUSCAN Indexes) suggested by guidance documents [41, 42], and, in fact, they proved to be quite reliable. The data are longitudinal in nature. Another study strength was that it describes the decision-making process leading to the selection of a single value from a range of different MCID cut offs by comparing the percentages of change scores exceeding the MCID. The process considered the major concordance between those based on anchor-based methods and those based on distribution-based approaches, privileging those based on the former [40, 43, 44], and evaluated the differences in terms of clinical OA.

## Conclusion

To conclude, the study shows that the AUSCAN hand and WOMAC hip/knee physical function scores are indeed sensitive to the effects of OA. The data analysed using various health and physical performance measures as external anchors showed that the minimally important decline over 1 year in the AUSCAN and WOMAC physical function scores was four and two points respectively. Further research is required to confirm the robustness of these estimates and to evaluate their temporal consistency and country-dependency.

## Abbreviation

AUSCAN: Australian/Canadian Hand Osteoarthritis hand Index; CI: Confidence Interval; EN: Edwards-Nunnally index; EN90: EN with 90% CI; EN95: EN with 95% CI; EPOSA: European Project on OsteoArthritis; EQ VAS: Health status using the visual analogue scale; EQ-5D-5 L: Health status using five dimensions; HADS: Hospital Anxiety and Depression Scales; IQR: interquartile range; MCID: Minimum Clinically Important Difference; ROC: Receiver



operating characteristic; SD: Standard deviation; SEM: Standard Error Measurement; SEM63: SEM with 63% CI; SEM90: SEM with 90% CI; SEM95: SEM with 95% CI; SRM: standardized response mean; SRM2: SRM with Cohen's threshold 0.20; SRM5: SRM with Cohen's threshold 0.50; SRM8: SRM with Cohen's threshold 0.80; WOMAC: Western Ontario and McMaster Universities knee and hip Index

### Acknowledgements

The authors would like to acknowledge the study's primary contributor, Thorsten Nikolaus, MD, a researcher who worked at the Bethesda Geriatric Clinic of the University of Ulm, Germany and died in September of 2013. We would also like to express our thanks to all of the individuals who participated in any and all ways in the EPOSA study.

Appreciation is also expressed to Linda Inverso Moretti for her assistance in editing the manuscript.

EPOSA Research Group

Nikolaus T, Peter R, Denkinge MD, Herbolsheimer F, Maggi S, Zambon S, Limongi F, Noale M, Siviero P, Deeg DJ, van der Pas S, Schaap LA, van Schoor NM, Timmermans EJ, Otero A, Castell MV, Sanchez-Martinez M, Quieipo R, Pedersen NL, Broumandi R, Dennison EM, Cooper C, Edwards MH, Parsons C.

### Funding

This work was supported by a non-commercial private funder.

The Indicators for Monitoring COPD and Asthma - Activity and Function in the Elderly in Ulm study (IMCA - ActiFE) was supported by the European Union [2005121] and the Ministry of Science, Baden-Württemberg. The Italian cohort study is part of the National Research Council Project on Aging (PNR). The Longitudinal Aging Study Amsterdam (LASA) is financially supported by the Dutch Ministry of Health, Welfare and Sports. The Peñagrande study was partially supported by the National Fund for Health Research (Fondo de Investigaciones en Salud) of Spain [FIS PI 05/1898; FIS RETICEF RD06/0013/1013 and FIS PS09/02143]. The Swedish Twin Registry is supported in part by the Swedish Ministry of Higher Education. The Hertfordshire Cohort Study is funded by the Medical Research Council of Great Britain, Arthritis Research UK, the British Heart Foundation and the International Osteoporosis Foundation.

### Availability of data and materials

The data that support the findings of this study are available from <https://www.eposa.org/> but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the EPOSA Research Group.

### Authors' contributions

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. PS had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study conception and design. SZ, CC, EMD, MHE, SvdP, LAS, NMvS, MDD, RP, MVC, AO, NLP, DJHD, SM. Acquisition of data. PS, FL, AG, SZ, CC, EMD, MHE, EJT, SvdP, LAS, NMvS, MDD, FH, RP, MVCI, AO, RQ, NLP, DJHD, SM. Analysis and interpretation of data. PS, FL, AG, SZ, CC, EMD, MHE, EJT, SvdP, LAS, NMvS, MDD, FH, RP, MVCI, AO, RQ, NLP, DJHD, SM.

### Ethics approval and consent to participate

All the participants gave written informed consent and the study design and procedures, outlined in detail elsewhere [4], were granted approval by local ethics committees (Germany: Universität Ulm Ethikkommission [312/08]. Italy: Comitato Etico Provinciale Treviso [XLIV-RSA/AULSS7]. The Netherlands: Medisch Ethische Toetsingscommissie Vrije Universiteit Amsterdam [2002/141]. Spain: Comité Ético de Investigación Clínica del Hospital Universitario La Paz Madrid [PI-1080]. Sweden: Till forskningsetikommittén vid Karolinska Institutet Stockholm [00-132]. UK: Hertfordshire Research Ethics Committee [10/H0311/59]).

### Consent for publication

Not applicable.

### Competing interests

CC: consultancy, lecture fees and honoraria from Alliance for Better Bone Health, Amgen, Eli Lilly, GSK, Medtronic, Merck, Novartis, Pfizer, Roche, Servier, Takeda, and UCB (less than \$ 10,000 each).

ED: speaking fees from Eli Lilly (less than \$ 10,000).

No competing interests were reported by the other authors.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>National Research Council, Neuroscience Institute - Aging Branch, Via Giustiniani 2, 35128 Padova, Italy. <sup>2</sup>Department of Medicine, University of Padova, Via 8 Febbraio 2, 35122 Padova, Italy. <sup>3</sup>MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton General Hospital, Tremona Road, Southampton SO16 6YD, UK. <sup>4</sup>Portsmouth Hospitals NHS Trust, Portsmouth PO6 3LY, UK. <sup>5</sup>Department of Epidemiology and Biostatistics, Amsterdam UMC - location VU University medical center, Amsterdam Public Health research institute, Amsterdam, the Netherlands. <sup>6</sup>Bethesda Geriatric Clinic, University of Ulm, Zollernring 26, 89073 Ulm, Germany. <sup>7</sup>Institute of the History Philosophy and Ethics of Medicine, University of Ulm, Frauensteige 6, 89075 Ulm, Germany. <sup>8</sup>Department of Preventive Medicine and Public Health, Unit of Primary Care and Family Medicine, Faculty of Medicine, Universidad Autonoma de Madrid, Arzobispo Morcillo 4, 28029 Madrid, Spain. <sup>9</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, P.O.Box 281, Nobels väg 12A, SE-171 77 Stockholm, Sweden.

Received: 16 January 2019 Accepted: 26 April 2019

Published online: 17 May 2019

### References

- Bellamy N, Campbell J, Haraoui B, Gerez-Simon E, Buchbinder R, Hobby K, et al. Clinimetric properties of the AUSCAN osteoarthritis hand index: an evaluation of reliability, validity and responsiveness. *Osteoarthritis Cartilage*. 2002; 10:863–9.
- 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.
- Roorda L, Jones C, Waltz M, Lankhorst G, Bouter L, van der Eijken JW, et al. Satisfactory cross cultural equivalence of the Dutch WOMAC in patients with hip osteoarthritis waiting for arthroplasty. *Ann Rheum Dis*. 2004;63:36–42.
- van der Pas S, Castell MV, Cooper C, Denkinge 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 Musculoskeletal Disord*. 2013;14:138–48.
- Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertain the minimal clinically important difference. *Control Clin Trials*. 1989;10:407–15.
- King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res*. 2011; 11:171–84.
- Bellamy N, Hochberg M, Tubach F, Martin-Mola E, Awada H, Bombardier C, et al. Development of multinational definitions of minimal clinically important improvement and patient acceptable symptomatic state in osteoarthritis. *Arthritis Care Res (Hoboken)*. 2015;67:972–80.
- Angst F, Aeschlimann A, Michel BA, Stucki G. Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. *J Rheumatol*. 2002;29:131–8.
- Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. *Arthritis Rheum*. 2001;45:384–91.
- Quintana JM, Escobar A, Bilbao A, Arostegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after hip joint replacement. *Osteoarthritis Cartilage*. 2005;13:1076–83.

11. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. *Ann Rheum Dis*. 2005;64:29–33.
12. Escobar A, Quintana JM, Bilbao A, Aróstegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. *Osteoarthr Cartil*. 2007;15:273–80.
13. Ornetti P, Dougados M, Paternotte S, Logeart I, Gossec L. Validation of a numerical rating scale to assess functional impairment in hip and knee osteoarthritis: comparison with the WOMAC function scale. *Ann Rheum Dis*. 2011;70:740–6.
14. Escobar A, García Pérez L, Herrera-Espiñeira C, Aizpuru F, Sarasqueta C, Gonzalez Sáenz de Tejada M, et al. Total knee replacement; minimal clinically important differences and responders. *Osteoarthr Cartil*. 2013;21:2006–12.
15. Clement ND, Bardgett M, Weir D, Holland J, Gerrand C, Deehan DJ. What is the minimum clinically important difference for the WOMAC index after TKA? *Clin Orthop Relat Res*. 2018;476(10):2005–14.
16. Zhang Y, Niu J, Kelly-Hayes M, Chaisson CE, Aliabadi P, Felson DT. Prevalence of symptomatic hand osteoarthritis and its impact on functional status among the elderly: the Framingham study. *Am J Epidemiol*. 2002;156:1021–7.
17. Dekker J, van Dijk GM, Veenhof C. Risk factors for functional decline in osteoarthritis of the hip or knee. *Curr Opin Rheumatol*. 2009;21:520–4.
18. Ostendorf M, van Stel HF, Buskens E, Schrijvers AJ, Marting LN, Verbout AJ, et al. Patient-reported outcome in total hip replacement: a comparison of five instruments of health status. *J Bone Joint Surg Br*. 2004;86:801–8.
19. Maly MR, Costigan PA, Olney SJ. Determinants of self-report outcome measures in people with knee osteoarthritis. *Arch Phys Med Rehabil*. 2006;87:96–104.
20. Bellamy N, Buchanan WW. A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee. *Clin Rheumatol*. 1986;5:231–41.
21. Barthel HR, Peniston JH, Clark MB, Gold MS, Altman RD. Correlation of pain relief with physical function in hand osteoarthritis: randomized controlled trial post hoc analysis. *Arthritis Res Ther*. 2010;12:R7.
22. Louie GH, Ward MM. Association of measured physical performance and demographic and health characteristics with self-reported physical function: implications for the interpretation of self-reported limitations. *Health Qual Life Outcomes*. 2010;8:84.
23. Altman RD. Hand Function in Osteoarthritis In: Duruöz MT, editor. *Hand function. A practical guide to assessment*. New York: Springer Science +Business Media; 2014. p. 63–9.
24. Liu R, Damman W, Kaptein AA, Rosendaal FR, Kloppenburg M. Coping styles and disability in patients with hand osteoarthritis. *Rheumatology (Oxford)*. 2016;55:411–8.
25. van Dijk GM, Veenhof C, Spreeuwenberg P, Coene N, Burger BJ, van Schaardenburg D, et al. Prognosis of limitations in activities in osteoarthritis of the hip or knee: a 3-year cohort study. *Arch Phys Med Rehabil*. 2010;91:58–66.
26. Roberts HC, Denison HJ, Martin HJ, Patel HP, Syddall H, Cooper C, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. *Age Ageing*. 2011;40:423–9.
27. Zigmond A, Snaith R. The hospital anxiety and depression scale. *Acta Psychiatr Scand*. 1983;67:361–70.
28. Brooks R, Rabin R, De CF. The measurement and valuation of health status using EQ-5D: a European perspective. Kluwer Academic Publishers; 2003. <https://www.springer.com/gp/book/9781402012143>.
29. Altman RD. Classification of disease: osteoarthritis. *Semin Arthritis Rheum*. 1991;20(Suppl 2):40–7.
30. Zhang W, Doherty M, Leeb BF, Alekseeva L, Arden NK, Bijlsma JW, et al. EULAR evidence-based recommendations for the diagnosis of hand osteoarthritis: report of a task force of ESCISIT. *Ann Rheum Dis*. 2009;68:8–17.
31. 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.
32. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61:102–9.
33. Perkins NJ, Schisterman EF. The Youden index and the optimal cut-point corrected for measurement error. *Biom J*. 2005;47:428–41.
34. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol*. 1995;48:1369–78.
35. Cohen J. *Statistical power analysis for the behavioral sciences*. New York: Academic Press; 1977.
36. Wywich 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;52:861–73.
37. Speer DC. Clinically significant change: Jacobson and Truax (1991) revisited. *J Consult Clin Psychol*. 1992;60:402–8 Erratum in: *J Consult Clin Psychol*. 1993;61:27.
38. Devji T, Guyatt GH, Lytvyn L, Brignardello-Petersen R, Foroutan F, Sadeghirad B, Buchbinder R, Poolman RW, Harris IA, Carrasco-Labra A, Siemieniuk RAC, Vandvik PO. Application of minimal important differences in degenerative knee disease outcomes: a systematic review and case study to inform BMJ rapid recommendations. *BMJ Open*. 2017;7(5):e015587. <https://doi.org/10.1136/bmjopen-2016-015587>.
39. Gandek B. Measurement properties of the Western Ontario and McMaster universities osteoarthritis index: a systematic review. *Arthritis Care Res (Hoboken)*. 2015;67:216–29.
40. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003;56:395–407.
41. Altman R, Brandt K, Hochberg M, Moskowitz R, Bellamy N, Bloch DA, et al. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the osteoarthritis research society. Results from a workshop. *Osteoarthr Cartil*. 1996;4:217–43.
42. Maheu E, Altman RD, Bloch DA, Doherty M, Hochberg M, Mannoni A, et al. Design and conduct of clinical trials in patients with osteoarthritis of the hand: recommendations from a task force of the osteoarthritis research society international. *Osteoarthr Cartil*. 2006;14:303–22.
43. Wells G, Beaton D, Shea B, Boers M, Simon L, Strand V, et al. Minimal clinically important differences: review of methods. *J Rheumatol*. 2001;28:406–12.
44. Wells G, Anderson J, Beaton D, Bellamy N, Boers M, Bombardier C, et al. Minimal clinically important difference module: summary, recommendations, and research agenda. *J Rheumatol*. 2001;28:452–4.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

