

What Preoperative Factors are Associated With Not Achieving a Minimum Clinically Important Difference After THA? Findings from an International Multicenter Study

Pakdee Rojanasopondist BA, Vincent P. Galea BA, James W. Connelly BA, Sean J. Matuszak BA, Ola Rolfson MD, PhD, Charles R. Bragdon PhD, Henrik Malchau MD, PhD

Received: 17 September 2018 / Accepted: 17 January 2019 / Published online: 15 May 2019
Copyright © 2019 by the Association of Bone and Joint Surgeons

Abstract

Background Despite innovations in THA, there remains a subgroup of patients who experience only modest pain relief and/or functional improvement after the procedure. Although several studies have previously sought to identify factors before surgery that were associated with achieving or not achieving a meaningful improvement after THA, there is no consensus on which factors are most associated;

The institution of one or more of the authors (PR, VPG, JWC, SJM, CRB, HM) has received, during the study period, research support funding in an amount of USD 100,001 to USD 1,000,000 from Zimmer Biomet.

Clinical Orthopaedics and Related Research® neither advocates nor endorses the use of any treatment, drug, or device. Readers are encouraged to always seek additional information, including FDA approval status, of any drug or device before clinical use. Each author certifies that his institution approved the human protocol for this investigation and that all investigations were conducted in conformity with ethical principles of research. Investigation performed at Harris Orthopaedics Laboratory, Massachusetts General Hospital, Boston, MA, USA.

P. Rojanasopondist, V. P. Galea, J. W. Connelly, S. J. Matuszak, C. R. Bragdon, H. Malchau, Harris Orthopaedics Laboratory, Massachusetts General Hospital, Boston, MA, USA

O. Rolfson, Department of Orthopaedics Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

P. Rojanasopondist (✉), Harris Orthopaedics Laboratory, Massachusetts General Hospital, 55 Fruit St. GRJ 1231, Boston, MA 02114, USA, Email: projanasopondist@mgh.harvard.edu

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

many studies have relied on single-center or single-country multicenter studies for their cohorts.

Questions/purposes We sought to identify (1) the proportion of patients who do not achieve a minimum clinically important difference (MCID) in pain and function 1 year after THA, and (2) the preoperative factors that were associated with not achieving MCIDs in pain and function 1 year after THA.

Methods This retrospective study analyzed data gathered from a prospective international, multicenter study examining the long-term clinical outcomes of two different polyethylene liners and two different acetabular shells. A total of 814 patients from 12 centers across four countries were enrolled in the study, with the final cohort consisting of 594 patients (73%) who all had complete preoperative and 1-year PROMs as well as a valid preoperative radiograph used to measure minimum joint space width. The outcomes in this study were achieving evidence-derived MCIDs in (1) pain, defined as a reduction of two points on an 11-point (0 = very little, 10 = worst imaginable) numerical rating scale (NRS) for hip-related pain or reporting a 1 year NRS-pain score of 0, and (2) function, defined as an increase equal to or greater than 8.3 on the SF-36 Physical Function subscore (range: 0 to 100; 0 = maximum disability, 100 = no disability) or reporting a 1-year SF-36 Physical Function subscore within the 95th percentile of scores in our cohort. All demographic variables, such as age, sex, country; surgical factors, including body mass index (BMI), surgical approach, acetabular liner type, and preoperative PROMs, were included as covariates in a binary logistic regression model. We used a backwards stepwise elimination algorithm to reach the simplest, best-fit model.

Results In the final analysis cohort of 594 patients, 54 patients (9%) did not achieve the MCID in pain and 146

(25%) patients did not achieve the MCID in physical function after THA. After controlling for potential confounding variables such as age, BMI, and preoperative PROMs, we found that higher joint space width (odds ratio (OR) = 2.19; 95% confidence interval (CI) = 1.49–3.22; $p < 0.001$), lower preoperative SF-36 Mental Component Summary (MCS) (OR = 0.95; 95% CI = 0.93–0.98; $p = 0.001$), and female sex (OR = 2.04; 95% CI = 1.08–3.82; $p = 0.027$) were associated with failing to achieve a MCID in pain. It is important to note that the effect size of having a higher preoperative SF-36 MCS is small, with a 1- or 10-point increase in SF-36 MCS decreasing the odds of a patient not achieving the pain MCID by 5% or 63%, respectively.

In a separate multivariable model, after controlling for potential confounding variables such as age, BMI, and preoperative PROMs, we found that higher joint space width (OR = 1.54; 95% CI = 1.18–2.02; $p = 0.002$), higher preoperative Harris hip score (HHS) (OR = 1.01; 95% CI = 1.00–1.03; $p = 0.019$) and undergoing surgery in Scandinavia (OR = 1.73; 95% CI = 1.17–2.55; $p = 0.006$) were associated with failing to achieve a MCID in physical function. It is important to note that the effect size of having a higher preoperative HHS is very small, with a 1- or 10-point increase in HHS increasing the odds of not achieving the physical function MCID by only 1% or 15%, respectively.

Conclusions These findings suggest that surgeons should counsel patients with high joint space width, female patients, and patients undergoing surgery in Scandinavia that they may be much less likely to experience meaningful pain relief or functional improvement after THA, and in light of that, determine whether indeed surgery should be postponed or avoided in those patients. Lower SF-36 MCS score and higher HHS before surgery were also found to be associated with not achieving MCIDs in pain and physical function, respectively, after surgery, but both had relatively small effect sizes. Future prospective studies may consider exploring the relationship between less pain relief or functional improvement and the risk factors identified in this study, such as high joint space width, to validate our findings and determine if the variables we identified are truly predictive of worse postoperative outcomes. Future retrospective studies of regional or national registry data should use the analysis methods presented within this study to both identify the portion of the THA patients who do not achieve a MCID in pain or physical function after surgery and confirm if the preoperative risk factors for poor improvement identified within our international, multicenter cohort are also found in a larger patient population with more diverse implants and comorbidities.

Level of Evidence Level III, therapeutic study.

Introduction

A subset of patients who undergo THA have persistent pain, unsatisfactory functional gains, and incomplete restoration of quality of life [3, 8]. A common approach to assessing if the basic aims of THA have been fulfilled is using patient-reported outcome measures (PROMs). As such, PROMs can help facilitate shared decision-making before a THA and can offer a meaningful way for patient-centered input to factor into value determination [5, 16, 17]. Despite their benefits, however, PROMs can be difficult to interpret because statistically significant but clinically irrelevant differences between groups can often be found [30]. One proposed method of correcting for this issue is by using a defined minimum clinically important difference (MCID) to assess improvement in a PROM after an intervention. The MCID is defined as the smallest improvement in a PROM that has been determined to be important to patients [26].

Previously, several studies have sought to identify preoperative factors that are associated with achieving meaningful improvements in pain or function after THA [7, 9–11, 18, 21, 24]. These studies have found factors correlated with postoperative PROMs to include age at surgery, body mass index (BMI), comorbidity burden, mental health, functional status, and radiographic osteoarthritis severity. There is, however, no clear consensus on which factors are most strongly associated with postoperative PROMs and, furthermore, only a few of these studies used a MCID threshold as an outcome [7, 9, 24]. Of the studies that used a MCID in a PROM as an outcome, all were either single-center studies or single-country multicenter studies [7, 24], which may limit their generalizability.

Therefore, we sought to identify (1) the proportion of patients who do not achieve a minimum clinically important difference (MCID) in pain and function 1 year after THA, and (2) the preoperative factors that were associated with not achieving MCIDs in pain and function 1 year after THA.

Patients and Methods

Study Design and Setting

This retrospective study analyzed data gathered from a prospective, international multicenter study evaluating the long-term clinical performance of two acetabular shells and two polyethylene liners from a single manufacturer. The data collected as part of this original study was well suited for the current retrospective study examining factors associated with postoperative improvement because the original study: (1) followed a cohort of patients who received a modern, widely-used implant system; (2) gathered data from a wide array of countries,

surgeons, and practice settings; and (3) collected a robust set of disease-specific and general health PROMs at both a preoperative and 1-year time point.

All patients underwent THA between 2007 and 2012 and received either a porous titanium-coated (Regenerex[®], Zimmer Biomet, Warsaw, IN, USA) or a plasma-sprayed (Ringloc[®], Zimmer Biomet) acetabular shell paired with either a vitamin E-diffused polyethylene (E1[®], Zimmer Biomet,) or a moderately-crosslinked polyethylene (ArComXL[®], Zimmer Biomet) acetabular liner. All patients also received a 32-mm or 36-mm ceramic or cobalt-chromium femoral head and a cementless Biomet femoral stem of the surgeon's choice. Of 814 enrolled study participants, 89 (11%) patients received a ceramic head and 24 (3%) patients received a 36-mm femoral head, and all patients who received either a ceramic head or a 36-mm head underwent surgery in Scandinavia. The three most common femoral stems implanted in the 814 enrolled study participants were the Taperloc (Zimmer Biomet, Warsaw, IN, USA) (*n* = 370, 46%), Bimetric HA (Zimmer Biomet, Warsaw, IN, USA) (*n* = 239, 29%), and Bimetric PC (Zimmer Biomet, Warsaw, IN, USA) (*n* = 162, 20%). Of the 370 patients who received Taperloc stems, all 370 underwent surgery within the United States. Of the 239 patients who received a Bimetric HA stem, one underwent surgery in the United States and 238 underwent surgery in Scandinavia. Of the 162 patients who received a Bimetric PC stem, one underwent surgery in the United States and 161 underwent surgery in Scandinavia. A total of 814 patients from 12 centers across the United States (*n* = 398) and Scandinavia (Denmark, *n* = 230; Norway, *n* = 51; and Sweden, *n* = 135) were enrolled into this study. All enrolled patients were between the ages of 20 and 75 years and diagnosed with primary osteoarthritis. Of 814 enrolled patients, 741 patients (91%) had complete preoperative PROMs, 683 patients (84%) had complete preoperative PROMs and a valid preoperative radiograph, and 594 patients (73%) had complete preoperative data and complete 1-year PROMs. Considering those with complete

preoperative data, there was no difference in age (*p* = 0.688) or sex (*p* = 0.080) between those with complete or incomplete 1-year data, but patients with incomplete 1-year data had a slightly higher average BMI (31 ± 7) than patients with complete 1-year data (29 ± 5) (*p* = 0.008). The final study cohort was representative of the population of patients eligible for a THA and demonstrated improvements in all PROMs (*p* < 0.001) between the preoperative and 1-year interval (Table 1) with the most pronounced improvements in disease-specific PROMs.

As part of the original study, all patients consented to be followed with plain radiographs and a set of PROMs preoperatively and at the 1-year interval (mean followup time: 1.1 years, range: 0.8–2.5 years). Demographic variables (age, sex, country), surgical data (procedure performed, surgical approach, and BMI), and component information (acetabular shell, acetabular liner, femoral head, and femoral stem) were reported for each patient at study enrollment. General health PROMs collected include the Short-Form 36 (SF-36), the EuroQol 5-dimension three-level (EQ-5D), and the University of California Los Angeles (UCLA) activity score. Disease-specific PROMs collected include the Harris hip score (HHS) and a numerical rating scale (NRS) for hip-related pain. All PROMs were administered on paper in the local language of each study site. After collection, all data were anonymized and transferred to an academic contract research organization (ACRO) at the Massachusetts General Hospital, Boston, USA, via a secure, web-based portal by all participating study sites. The original study protocol was approved by the institutional review board at each respective study site and at the ACRO. Additional consent was not required for the present analysis.

Outcomes

The SF-36 is a generic, health-related quality of life metric. Its 36 items cover eight domains: physical function, role physical, bodily pain, general health, vitality, social function,

Table 1. Preoperative, 1-year PROM values and median PROM change for all patients in final analysis cohorts

PROM	Preoperative value (<i>n</i> = 594)	1-year value (<i>n</i> = 594)	PROM change (<i>n</i> = 594)	<i>p</i> value
EQ-Index	0.6 (0.2-0.7)	0.9 (0.7-1.0)	0.3 (0.1-0.6)	< 0.001
EQ-VAS	70 (54-82)	88 (80-94)	12 (2-26)	< 0.001
SF36-MCS	53 (44-60)	58 (52-61)	2.6 (-3-11)	< 0.001
SF36-PCS	37 (31-42)	53 (45-57)	13 (8-21)	< 0.001
HHS	52 (39-64)	96 (86-100)	39 (27-50)	< 0.001
NRS-Pain	6 (4-7)	1 (0-2)	-5 (-6-3)	< 0.001
UCLA	5 (4-7)	6 (5-7)	1 (0-3)	< 0.001

All values presented as median (IQR); groups compared using a Mann-Whitney U test; PROM = patient-reported outcome measure; VAS = visual analog scale; MCS = mental component summary; PCS = physical component summary; HHS = Harris hip score; NRS = numerical rating scale; UCLA = University of California Los Angeles Activity score.

role emotional, and mental health. Each domain has a calculated subscore that ranges from 0 to 100, with 0 indicating maximum disability and 100 indicating no disability, and also contributes to two summary scores: the Physical Component Summary (PCS) and the Mental Component Summary (MCS) [41].

Patients also completed an 11-point NRS-Pain metric in response to the question “Indicate your average pain due to your most recently diagnosed/treated hip during the past month.” Possible responses ranged from 0 (very little) to 10 (worst imaginable).

Covariates

The EQ-5D is a generic, health-related quality of life instrument that evaluates patients in five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension is divided into three levels and can be summarized as a global health index with a weighted total value (British value set used) ranging from -0.594 to a maximum of 1.0 [12, 13]. Additionally, the EQ-5D also includes a VAS (EQ-VAS) used to represent general health state that ranges from 0 (worst imaginable) to 100 (best imaginable).

The UCLA Activity Score is an instrument that classifies patients' physical activity on a 0 to 10 scale, with 10 corresponding to a high level of physical activity [33].

The HHS is a joint-specific PROM often used to evaluate improvements after THA. The HHS is a composite measure covering pain and physical function and is scored on a scale ranging from 0 to 100, with a higher score representing improved function and decreased pain [20, 37].

Preoperative AP pelvic radiographs, with the patient lying supine and feet with an internal rotation of 10°, from within 6 months before the index surgery were also collected and used to calculate the patient's minimum joint space width, the shortest distance between the femoral head margin and the acetabulum. Digital measurements were calibrated by using a calibration ball of known size. If unavailable, we calibrated measurements by using the known size of the patient's replaced femoral head in the postoperative film to measure the size of the contralateral femoral head, which was then used to calibrate the preoperative film. One-year AP pelvic radiographs were also collected and used to measure acetabular cup positioning using Hip Analysis Suite software (University of Chicago, Chicago, IL, USA). All joint space width and cup positioning readings were completed by a single, board-certified orthopaedic surgeon (KM), who was blinded to patient demographics and clinical outcomes. All joint space width measurements were completed using the mDesk™ software package (RSA Biomedical, Umeå, Sweden), recorded in millimeters, and treated as a continuous variable in statistical

modeling. We assessed intrareader variability for measuring joint space width by comparing two sets of joint space width measurements of 10 randomly selected films from each study site, with the readings done 3 weeks apart to minimize interpretation bias and calculating a kappa value. Reliability for joint space width analysis proved to be excellent ($\kappa = 0.9$).

Statistical Analysis, Study Size

The outcomes considered in this study were achieving a MCID in physical function as defined by the systematic review of Maltenfort et al. [30], and achieving a MCID in pain as defined by Farrar et al. [15], a research article, at 1 year after THA. We chose these two studies because they, respectively, reviewed studies that used an anchor-based method or employed an anchor-based approach to generate a MCID value. The anchor-based approach is preferred because it associates a numerical change in a PROM instrument to a patient-reported assessment of improvement [32]. To achieve an MCID in function, a patient had to achieve an increase equal to or greater than 8.3 on the SF-36 Physical Function subscore or report a 1-year SF-36 Physical Function subscore of equal to or greater than 57.2, the cutoff for being within the 95th percentile of scores of our cohort [30]. To achieve an MCID in pain, a patient had to achieve either a reduction of two points on the NRS-Pain or report a 1-year NRS-Pain score of 0 [15].

We conducted a post hoc power analysis to determine the number of patients needed to detect a 5% or 10% difference in the proportion of patients who did not achieve an MCID in either pain or physical function with 80% power and a Type I error rate of 0.05. Assuming that approximately 10% of patients would not achieve the MCID in pain [8], to detect a 5% or 10% difference in the proportion of patients who would not achieve the MCID in pain, a minimum of 682 or 196 patients was needed, respectively. Assuming that approximately 20% of patients would not achieve the MCID in physical function [29], to detect a 5% or 10% difference in the proportion of patients who would not achieve the MCID in physical function, a minimum of 1089 or 259 patients was needed, respectively.

All variables were entered into a multivariable binary logistic regression predicting failure to achieve a MCID in either pain or physical function. Variables tested included: demographic and surgical factors, general and mental health state, implant variables, and preoperative joint space width (Table 2). Preoperative NRS-Pain and SF-36 PROM scores were excluded from models testing failure to achieve a MCID in pain and a MCID in physical function, respectively, to prevent overfitting. All other preoperative PROMs were included in both models. Femoral head material, femoral head size, and femoral stem type were also

Table 2. Demographics, preoperative PROM values, and multivariable model results for patients who did and did not achieve a minimum clinically important difference in pain

Variable	All (n = 594)	Achieved pain MCID (n = 540)	Did not achieve pain MCID (n = 54)	Multivariable p values (β; odds ratio; 95% CI)
Age (years)*	60 (26-75)	61 (26-75)	60 (36-75)	–
Female (vs male) sex [†]	268 (45)	246 (46)	22 (41)	0.027 (0.71; 2.04; 1.08-3.82)
BMI (kg/m ²)*	29 (17-46)	29 (17-46)	27 (21-36)	–
Scandinavian (vs US) [†]	307 (52)	273 (51)	34 (63)	0.056 (-0.59; 1.80; 0.98-3.28)
Posterolateral (vs anterolateral) approach [†]	346 (58)	309 (57)	37 (69)	–
VEPE (vs ModXL) liner [†]	339 (57)	233 (43)	26 (48)	0.070 (0.57; 1.77; 0.96-3.27)
PTC shell (vs PS shell) [†]	259 (44)	303 (56)	36 (67)	–
Abduction (°)*	42 (24-64)	42 (24-64)	44 (29-59)	–
Anteversion (°)*	16 (0-40)	16 (0-40)	16 (0-36)	–
Joint space width (mm)*	0.7 (0-4)	0.7 (0-4)	1 (0-3)	< 0.001 (0.78; 2.19; 1.49-3.22)
Preoperative EQ-Index [‡]	0.6 (0.2-0.7)	0.6 (0.2-0.7)	0.6 (0.2-0.7)	–
Preoperative EQ-VAS [‡]	70 (54-82)	70 (54-84)	70 (53-80)	–
Preoperative SF36-MCS [‡]	53 (44-60)	54 (45-61)	46 (39-57)	0.001 (-0.05; 0.95; 0.93-0.98)
Preoperative SF36-PCS [‡]	37 (31-42)	37 (31-42)	38 (31-43)	–
Preoperative HHS [‡]	52 (39-64)	52 (39-63)	54 (40-69)	–
Preoperative NRS-Pain [‡]	6 (4-7)	6 (5-7)	4 (2-6)	– [§]
Preoperative UCLA [‡]	5 (4-7)	5 (4-7)	5 (4-7)	–

*Values presented as mean (range).

†values presented as n (%).

‡values presented as median (interquartile range [IQR]).

§variable excluded from multivariable, binary logistic regression model to prevent overfitting; MCID = minimum clinically important difference; BMI = body mass index; VEPE = vitamin-E highly cross-linked (130kGy) polyethylene; ModXL = moderately cross-linked (50kGy) and mechanically annealed polyethylene; PTC = porous titanium coated; PS: plasma sprayed; MCS = mental component summary; PCS = physical component summary; HHS = Harris hip score; NRS = numerical rating scale; UCLA = University of California Los Angeles Activity score.

excluded from models testing failure to achieve an MCID in pain and an MCID in physical function because all three variables were strongly collinear with country (United States versus Scandinavia). For the remaining variables, a backwards stepwise elimination algorithm was used to reach the simplest, best-fit model. To assess how well the final predictive model acted as a predictor, receiver operating curve (ROC) analysis was used to plot predictive model probabilities against whether or not a patient achieved an MCID in pain or an MCID in physical function.

Lastly, to test if achieving an MCID in pain was correlated to achieving an MCID in physical function, we conducted a chi-square test between the variables. Significance was set at $p < 0.05$ for all comparisons. The SPSS Statistics Version 24.0 (IBM, Armonk, New York, USA) software package was used for all analyses.

Results

Of the final cohort, 54 (9%) of 594 patients did not achieve the MCID in pain (Fig. 1). In addition, 146 patients (25%)

did not achieve the MCID in physical function after THA (Fig. 2).

After controlling for confounding variables, such as age, BMI, acetabular shell and liner type, and preoperative PROMs (Table 2), we found that greater joint space width (odds ratio [OR], 2.19; 95% confidence interval [CI] = 1.49-3.22; $p < 0.001$), lower preoperative SF-36 Mental component summary (MCS) (OR, 0.95; 95% CI = 0.929–0.980; $p = 0.001$), and female sex (OR, 2.04; 95% CI = 1.08–3.82; $p = 0.027$) were independently associated with not achieving a MCID in pain (Fig. 3). It is important to note that the effect size of having a lower preoperative SF-36 MCS is small, with a 1- or 10-point decrease in SF-36 MCS raising the odds of a patient not achieving the pain MCID by 5% or 63%, respectively. ROC analysis to assess the predictive capability of the MCID Pain model yielded an area under the curve (AUC) of 0.73 (95% CI = 0.66–0.80; $p < 0.001$).

In a separate model, after controlling for potential confounding variables such as age, BMI, acetabular shell and liner type, and preoperative PROMs (Table 3), we found that higher joint space width (OR, 1.54; 95% CI,

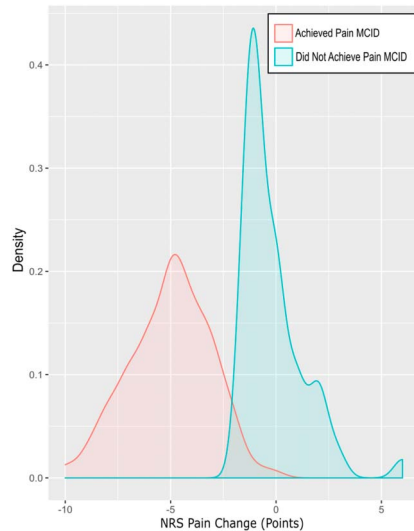


Fig. 1 Density plot shows the NRS Pain change (1-year minus preoperative) split by patients who achieved the Pain MCID and patients who did not achieve the Pain MCID.

1.18–2.02; $p = 0.002$), higher preoperative HHS (OR, 1.01; 95% CI, 1.00–1.03; $p = 0.019$), and undergoing surgery in Scandinavia (OR, 1.73; 95% CI, 1.17–2.55; $p = 0.006$) were found to be independently associated with not achieving an MCID in physical function (Fig. 4). It is important to note that the effect size of having a higher preoperative HHS is very small, with a 1- or 10-point increase in HHS increasing the odds of not achieving the physical function MCID by only 1% or 15%, respectively. ROC analysis to assess the predictive capability of the MCID

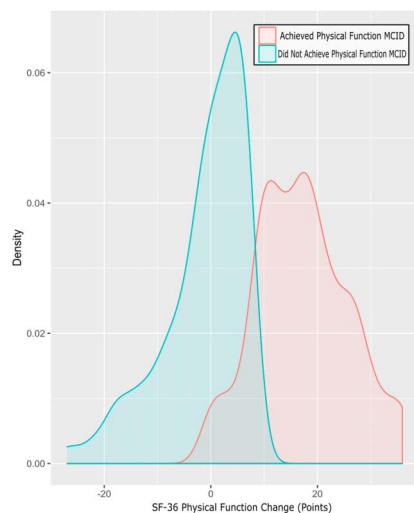


Fig. 2 Density plot shows the SF-36 Physical Function change (1-year minus preoperative) split by patients who achieved the Physical Function MCID and patients who did not achieve the Physical Function MCID.

Physical Function model yielded an AUC of 0.63 (95% CI = 0.58–0.69; $p < 0.001$).

Lastly, achieving an MCID in pain was found to be associated with achieving an MCID in Physical Function ($p < 0.001$).

Discussion

More than 300,000 THAs are performed each year in the United States, making the procedure one of the most commonly performed elective orthopaedic surgeries [1]. Although most patients experience decreased pain and improved physical function after recovery from surgery, some do not [8]. To better identify those patients likely and unlikely to benefit from major elective surgery, we sought to pinpoint the factors associated with meaningful improvements in pain and function after THA by including all variables in a binary logistic regression model and eliminating variables that did not add any value to the model via a backwards stepwise elimination algorithm. This model permits the analysis of all variables and does not exclude those that may not achieve statistical significance in univariate tests but may do so once other variables are considered. We found that higher joint space width, female sex, and poor SF-36 MCS were associated with not achieving an MCID in pain, and we also found that higher joint space width, higher preoperative HHS, and undergoing surgery in Scandinavia were associated with not achieving an MCID in physical function.

This study is not without limitations. First, only supine, AP plain radiographs were available to evaluate the severity of each patient's preoperative OA. Neither alternate preoperative radiograph views, such as a weightbearing AP radiograph or a shoot-through lateral, nor additional information on the length of preoperative OA symptom duration were available to provide a more complete picture of each patient's preoperative disease state. Despite this limitation, however, one study has found little difference between joint space width measured from supine radiographs and from weightbearing radiographs [4], and several other peer-reviewed publications exploring preoperative joint space width and its effects on postoperative THA PROMs have similarly relied upon supine AP radiographs [2, 38, 39].

Second, the value of an MCID and the resulting study outcomes can vary based on the method used to calculate the MCID. In this study, we chose to use an MCID for pain change as defined in a study by Farrar et al. [15], which calculated an MCID for a similar 11-point NRS (0 = no pain, 10 = worst possible pain) using a validated, 7-point categorical scale measuring the patient's global impression of change (1 = very much improved, 7 = very much worse) as an anchor. The data analyzed was collected as part of

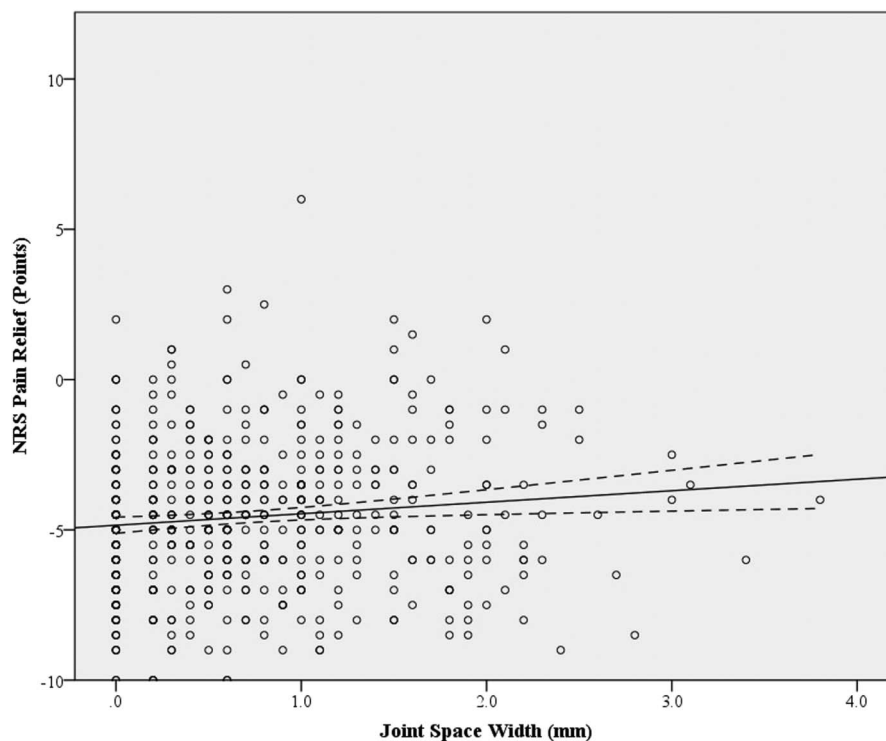


Fig. 3 Scatterplot displays NRS Pain relief against preoperative joint space width. The dashed line represents the 95% confidence interval for the linear best-fit line. The linear equation for the best-fit line is $y = 0.38x - 4.85$ (R-squared = 0.012, 95% confidence interval = 0.10–0.67) and demonstrates the relationship between increasing joint space width and decreased pain relief after THA.

10 double-blind, placebo-controlled, parallel, multicenter chronic pain studies. Although this study did not generate an MCID specifically for a population of patients with end-stage osteoarthritis undergoing THA, we chose this study because (1) to the best of the authors' knowledge there does not exist another study defining a MCID for a NRS-Pain instrument for patients undergoing THA, (2) it combined data from 10 multicenter studies with a total enrollment of 2879 patients, which covered a wide range of indications in both neuropathic and non-neuropathic chronic pain, (3) it used an anchor-based approach to developing its MCID, and (4) its primary findings were consistent with another published study [14] of acute breakthrough cancer pain which similarly found that a change score of -2.0 on a NRS-Pain scale was associated with the study's clinically important outcome of a patient's need to take additional pain medication. As for the MCID in physical function change, we chose to use an MCID identified in a systematic review performed on PubMed in September 2016 by Maltenfort et al. [30]. We chose to use the MCID value for the SF-36 Physical Function subscore identified by this study because this MCID was not only developed for a patient population with osteoarthritis undergoing a primary hip replacement, but also because the review focused primarily on articles

that used an anchor-based method. In both instances, we chose to apply evidence-defined MCIDs rather than calculate our own using distribution based methods, which would fail to link the numeric changes of a PROM to any kind of measurement of what is actually meaningful to a patient [26, 30, 32].

Third, no formal, a priori power analysis was conducted before beginning the study to determine the number of patients who needed to be enrolled to detect a 5% or 10% difference between groups in the proportion of patients who did not achieve an MCID in pain or physical function after THA. However, a post hoc power analysis revealed that our study contained enough patients to detect at least a 10% difference between groups in the proportion of patients who did not achieve an MCID in either pain or physical function.

Fourth, the fact that this study had a stricter inclusion criteria of requiring both a valid preoperative radiograph as well as complete PROMs, instead of requiring only complete PROMs as required by the original study, resulted in an initial cohort of only 683 (84%) out of 814 patients with complete preoperative data and, eventually, a final cohort of only 594 (73%) out of 814 patients with complete preoperative data and 1-year PROMs. Although this final

Table 3. Demographics, preoperative PROM values, and multivariable model results for patients who did and did not achieve an MCID in physical function

Variable	All (n = 594)	Achieved function MCID (n = 448)	Did not achieve function MCID (n = 146)	Multivariable p values (β; odds ratio; 95% CI)
Age (years)*	60 (26-75)	60 (32-75)	61 (26-75)	–
Female (vs male) sex [†]	268 (45)	197 (44)	71 (49)	–
BMI (kg/m ²)*	28.5 (17.4-46.2)	28.5 (17.4-46.0)	28.3 (19.5-46.2)	–
Scandinavian (vs US) [†]	307 (52)	216 (48)	91 (62)	0.006 (-0.55; 1.73; 1.17-2.55)
Posterolateral (vs anterolateral) approach [†]	346 (58)	253 (56)	93 (64)	–
VEPE (vs ModXL) liner [‡]	339 (57)	200 (45)	59 (40)	–
PTC shell (vs PS shell) [‡]	259 (44)	256 (57)	83 (57)	–
Abduction (°)*	42 (24-64)	42 (24-64)	43 (27-59)	–
Anteversion (°)*	16 (0-40)	16 (0-40)	16 (0-36)	–
Joint space width (mm)*	0.7 (0-4)	0.6 (0-4)	0.8 (0-3)	0.002 (0.43; 1.54; 1.18-2.02)
Preoperative EQ-Index [‡]	0.6 (0.2-0.7)	0.6 (0.2-0.7)	0.6 (0.2-0.7)	–
Preoperative EQ-VAS [‡]	70 (54-82)	70 (54-84)	69 (56-80)	–
Preoperative SF36-MCS [‡]	53 (44-60)	54 (44-60)	53 (44-61)	–
Preoperative SF36-PCS [‡]	37 (31-42)	36 (31-41)	39 (34-43)	– [§]
Preoperative HHS [‡]	52 (39-64)	51 (38-62)	54 (42-68)	0.019 (0.01; 1.01; 1.00-1.03)
Preoperative NRS-Pain [‡]	6 (4-7)	6 (5-7)	5 (4-7)	–
Preoperative UCLA [‡]	5 (4-7)	5 (3-6)	6 (4-7)	–

*Values presented as mean (range).

†values presented as n (%).

‡values presented as median (IQR).

§variable excluded from multivariable, binary logistic regression model to prevent overfitting; MCID = minimum clinically important difference; BMI = body mass index; VEPE: vitamin-E highly cross-linked (130kGy) polyethylene; ModXL: moderately cross-linked (50kGy) and mechanically annealed polyethylene; PTC: porous titanium coated; PS: plasma sprayed; MCS = mental component summary; PCS = physical component summary; HHS = Harris hip score; NRS = Numerical Rating scale; UCLA = University of California Los Angeles Activity score.

analysis cohort fell short of having at least an 80% followup percentage, we found no difference in age or sex between patients with complete preoperative data but incomplete 1-year data and patients with complete data at both intervals. Although we found that patients with complete preoperative data and incomplete 1-year data had higher BMI than patients with complete data, we did not find this difference in BMI to be significant in the final multivariable models for not achieving either the pain or physical function MCID.

Fifth, this study was unable to control for all factors that might influence the level of pain relief or physical function improvement experienced by study participants such as the comorbidity burden at the time of surgery or experience level of the surgeon. Despite this limitation, the addition of comorbidity data to our model would have likely not changed our results because one criteria for inclusion within the original study was the absence of any previous infection, osteoporosis, metabolic disorders that may impair bone formation, or any other major medical complications that could limit their ability to return for followup

for up to 10 years, which would have excluded many patients with serious comorbidities that could have negatively influenced their postoperative PROMs. Additionally, there was not much variation in surgeon experience within our study cohort as all participating surgeons in this original study were the equivalent of either attending or consultant orthopaedic surgeons specializing in adult hip reconstruction at their respective sites.

Lastly, the use of a MCID to analyze postoperative PROMs can be limited by the instrument's dichotomization of a continuous instrument, which can lead to a loss of information, as well as its susceptibility to being correlated to baseline PROM values. In instances where PROMs are subject to a ceiling effect, this limitation can result in an artificially low number of patients achieving an MCID after surgery. In this study, we attempted to control for this limitation by considering patients who reported no hip-related pain at 1-year as having achieved the MCID in pain and by considering patients with a 1-year SF-36 physical function subscore within the top 95th percentile of our cohort as having achieved the MCID in physical function.

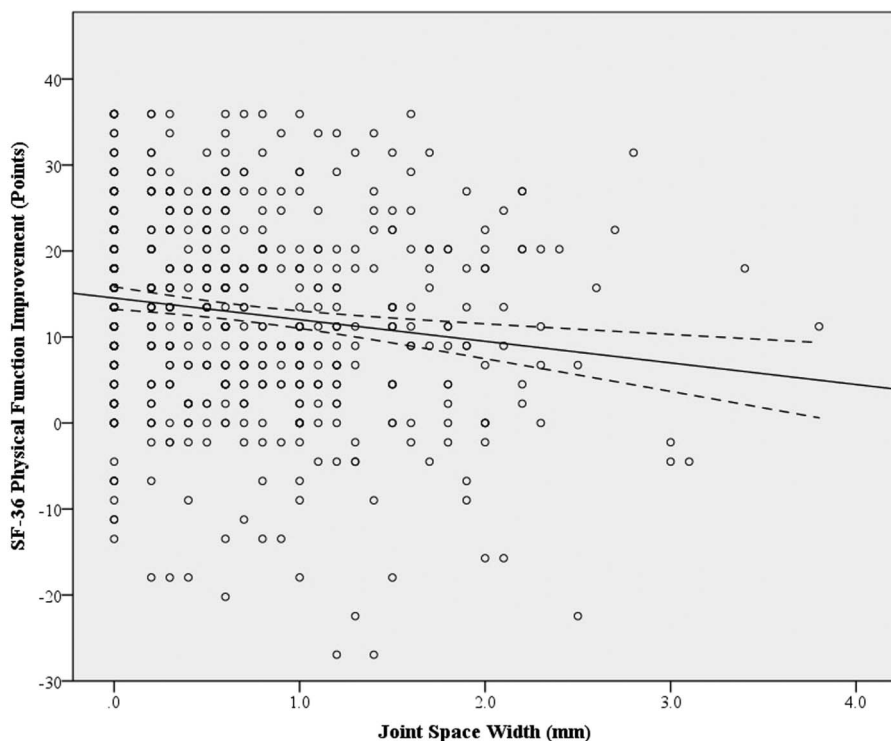


Fig. 4 Scatterplot displays SF-36 Physical Function improvement against preoperative joint space width. The dashed line represents the 95% confidence interval for the linear best-fit line. The linear equation for the best-fit line is $y = -2.51x + 14.53$ (R-squared = 0.021, 95% Confidence Interval = -3.89 to -1.14) and demonstrates the relationship between increasing joint space width and less physical function improvement after THA.

Using an international, multicenter cohort that received modern, widely-used implants and a robust battery of both disease-specific and general health PROMs, this study found that 54 (9%) out of 594 patients did not achieve the MCID in pain and 146 (25%) out of 594 patients did not achieve the MCID in physical function 1 year after THA. These results are in line with both the high rate of success of primary THAs and the findings of previous studies, which have reported that 7% to 23% of patients do not experience a meaningful decrease in pain [8], and approximately 20% of patients do not experience meaningful improvements in function after THA [29]. These results highlight that, despite the major technological and surgical advancements that have made THA as successful as it is today, there remains substantial room for improvement in how surgeons identify which patients are likely or unlikely to benefit from receiving a THA. Although the current analysis incorporates data from multiple centers, it is essential for future studies to confirm if our findings remain consistent within a population of patients who received a broader array of implants, had a more diverse comorbidity burden, and underwent surgery within more varied practice settings – a study topic that national or regional registries are uniquely positioned to answer. Given commonly cited

projections that THA demand is estimated to grow by 174% by 2030 [25], the need for future registry studies to identify how many and which patients experience either a decline or suboptimal improvements in pain or physical function after THA will be critical as the number of patients who fall into this category will likely only grow without targeted study and intervention.

Furthermore, we found that patients with higher joint space width, lower preoperative SF-36 MCS, and female patients were less likely to achieve an MCID in pain, and we found that patients with higher joint space width, higher preoperative HHS, and those who underwent surgery in Scandinavia were less likely to achieve an MCID in physical function. From the present analysis, we found that each one millimeter increase in joint space width increased the odds of not achieving the MCID in pain by 119% and increased the odds of not achieving the MCID in physical function by 54%. This relationship is consistent with previous, single-center studies [38, 39], single-country multicenter studies [23, 24, 40], and a systematic review [21] that have all reported that patients with less severe osteoarthritis, as measured by continuous joint space width or by the Kellgren-Lawrence scale, have worse functional improvement, pain relief, and lower activity level after THA

than patients with more severe osteoarthritis [38, 39]. Our findings on joint space width, combined with previously published results, are important as they provide evidence that multiple patients within our analysis cohort may have undergone the procedure prematurely. Given that THA is a major elective surgical procedure with multiple risks and high cost, it is absolutely critical that surgeons not only counsel patients with less radiographically severe osteoarthritis on their significantly increased odds of experiencing only a minor increase, if not a decrease, in their postoperative PROMs, but also provide the patient with alternative treatment options, such as delaying surgery or pursuing more conservative treatment options until their disease progresses further. It should be noted, however, that two single-center studies [2, 34] have previously reported finding no relationship between preoperative radiographic osteoarthritis severity and postoperative PROMs, and that these studies highlight the multifactorial relationship that links preoperative radiographic osteoarthritis status and postoperative PROMs.

In addition to joint space width, the present analysis found that a 1- or 10-point decrease in SF-36 MCS increased the odds of a patient not achieving the pain MCID by 5% or 63%, respectively. These findings are consistent with previous evidence that has demonstrated a link between poor emotional health and decreased pain relief after THA [6, 7, 35, 36]. Although the effect size of having a lower SF-36 MCS is relatively small in this analysis, as reported by Ayers et al. [6], patients with poor preoperative MCS scores may be more likely to use catastrophizing coping mechanisms and have a more difficult time with pain control after surgery. Therefore, surgeons may still want to consider discussing the option of delaying surgery with patients who report a low preoperative SF-36 MCS, which would allow them to receive counselling to improve their mental and emotional health before surgery.

Female sex was also found to increase the odds of a patient not achieving the pain MCID by 104% as compared with male sex. This finding is consistent with previous studies that have found that females are more likely to experience persistent postoperative pain after THA [27, 28]. Consequently, surgeons may counsel their female patients on their increased odds for not achieving an MCID in pain after THA to help manage their postoperative expectations. The relationship between female sex and postoperative PROMs, however, is mixed and several studies have reported no relationship between female sex and worse postoperative pain, function, or health-related quality of life [21, 31, 34].

This study found that a 1- or 10-point increase in preoperative HHS increased the odds of a patient not achieving the physical function MCID by 1% or 15%, respectively. Our finding that higher HHS was associated with not achieving the MCID in physical function is

consistent with previous evidence that has found a link between high preoperative function and decreased postoperative functional improvement [22, 23, 29]. However, it is important to note that the effect size of the association that we found between high preoperative HHS and not achieving an MCID in physical function is small. Although patients presenting with a high HHS should not be as worrisome as those with high preoperative joint space width, for example, surgeons should still consider discussing the slightly increased chance that patients with high HHS have of experiencing only minor postoperative improvements in function.

Lastly, this analysis also found that undergoing surgery in Scandinavia as opposed to in the United States increased the odds of not achieving the physical function MCID by 73%. Our finding that undergoing surgery in a Scandinavian country was associated with not achieving an MCID in physical function has not been previously reported. To explore this finding, we first compared the preoperative characteristics of US and Scandinavian patients in our final cohorts using univariate tests. We found that there was no difference between the two groups in age ($p = 0.102$), sex ($p = 0.284$), joint space width ($p = 0.257$), SF-36 MCS ($p = 0.997$), or preoperative HHS ($p = 0.339$). We found that US patients had higher BMIs ($p < 0.001$), preoperative NRS-Pain scores ($p < 0.001$), and preoperative EQ-5D Health-state ($p < 0.001$) than Scandinavian patients. We also found that Scandinavian patients had higher preoperative UCLA scores ($p = 0.032$), preoperative EQ-5D index ($p = 0.007$), and preoperative SF-36 PCS ($p = 0.004$) than US patients. Then, we repeated the full binary regression model for not achieving an MCID in physical function with interaction terms for country (Scandinavia versus US) and age, sex, BMI, and all PROMs (excluding SF-36 PCS to prevent overfitting) that were different in univariate testing. We repeated our model after adding these seven interaction terms and found that none of them added any value to the final model, and we subsequently excluded them. Lastly, because a SF-36 PCS interaction term was excluded to prevent overfitting, we performed additional univariate tests for the variable, and we found no difference in SF-36 PCS change ($p = 0.057$) or 1-year SF-36 PCS scores ($p = 0.480$) between Scandinavian and US patients. Therefore, we theorized that the differences observed in this study may be attributable to either cultural differences between how Scandinavian and American patients respond to PROMs, with Scandinavian patients more likely to report higher physical function before surgery, or differences in patient selection, with Scandinavian patients more likely to receive a THA at lower levels of functional impairment [19].

Our findings highlight the need for retrospective registry studies to examine if the risk factors we found for not achieving MCIDs in pain or function after THA are similar to those found in regional or country-wide cohorts of THA

patients. Alongside registry studies, future prospective studies should be conducted to determine if higher joint space width, lower SF-36 MCS, female sex, undergoing surgery in Scandinavia, and higher HHS before surgery are truly predictive of increased odds of not achieving meaningful improvements in pain or physical function. Lastly, both retrospective registry studies and prospective studies should be designed to explore the possible differences that may exist in how patients from different countries respond to commonly collected PROM instruments to help analyze and compare PROM data collected across different regions or countries.

In conclusion, we found that patients with higher joint space width, female sex, and undergoing surgery in Scandinavia were much less likely to achieve an MCID in either pain or physical function 1 year after THA, and that patients with low MCS or high HHS before surgery were slightly less likely to see improvements in pain and function. Based on these findings, surgeons should pause before performing a THA on patients who present with high joint space width or other risk factors for poor postoperative improvement. They should provide the patient with their increased odds of not achieving levels of pain reduction or physical function improvements that other patients have found to be meaningful after THA. Surgeons may also use this data to encourage patients to undergo alternative preoperative treatments if they have modifiable risk factors for not achieving meaningful improvements in pain or function after surgery. We emphasize, however, that the results of this study should not be used as appropriateness criteria for undergoing THA. Rather, it should be used as a tool that can guide the conversations that an orthopaedic surgeon may have with their patients, one that should empower patients by helping them understand their likelihood for improvement given their preoperative status. THA is a major surgical procedure that carries serious risks, and patients electing to undergo the procedure deserve more than a minor decrease in pain or a small increase in physical function in exchange for the risk they accept and the time and cost they pay.

Future retrospective studies of regional or national registry data should use the analysis methods presented within this study to both identify the portion of the THA patients who do not achieve an MCID in pain or physical function after surgery and confirm if the preoperative risk factors for poor improvement identified within our international, multicenter cohort are also found in a larger patient population with more diverse implants and comorbidities. Furthermore, future prospective studies should explore the relationship between less pain relief or functional improvement and the risk factors identified in this study to determine if the variables we identified are truly predictive of worse postoperative outcomes. Such analyses will be important for identifying ways to meaningfully interpret PROMs before

and after THA, and they will help to continue improving the outcomes and value of THA procedures.

Acknowledgments We thank Kirill Gromov MD, PhD, for his assistance in performing all joint space width and acetabular cup positioning measurements used within this manuscript.

References

1. Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project (HCUP). Available at: <http://hcupnet.ahrq.gov/>. Accessed October 9, 2018.
2. Al-Amiry BS, Gaber JF, Kadum BK, Brismar TB, Sayed-Noor AS. The influence of radiological severity and symptom duration of osteoarthritis on postoperative outcome after total hip arthroplasty: a prospective cohort study. *J Arthroplasty*. 2018;33:436–440.
3. Anakwe RE, Jenkins PJ, Moran M. Predicting dissatisfaction after total hip arthroplasty: a study of 850 patients. *J Arthroplasty*. 2011;26:209–13.
4. Auleley GR, Rousselin B, Ayral X, Edouard-Noel R, Dougados M, Ravaud P. Osteoarthritis of the hip: agreement between joint space width measurements on standing and supine conventional radiographs. *Ann Rheum Dis*. 1998;57:519–23.
5. Ayers DC, Bozic KJ. The importance of outcome measurement in orthopaedics. *Clin Orthop Relat Res*. 2013;471:3409–11.
6. Ayers DC, Franklin PD, Trief PM, Ploutz-Snyder R, Freund D. Psychological attributes of preoperative total joint replacement patients: implications for optimal physical outcome. *J Arthroplasty*. 2004;19:125–30.
7. Berliner JL, Brodke DJ, Chan V, SooHoo NF, Bozic KJ. John Charnley Award: Preoperative patient-reported outcome measures predict clinically meaningful improvement in function after THA. *Clin Orthop Relat Res*. 2016;474:321–329.
8. 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:e000435.
9. Clement ND, MacDonald D, Howie CR, Biant LC. The outcome of primary total hip and knee arthroplasty in patients aged 80 years or more. *J Bone Joint Surg Br*. 2011;93:1265–70.
10. Cushnaghan J, Coggon D, Reading I, Croft P, Byng P, Cox K, Dieppe P, Cooper C. Long-term outcome following total hip arthroplasty: a controlled longitudinal study. *Arthritis Rheum*. 2007;57:1375–80.
11. Davis AM, Wood AM, Keenan ACM, Brenkel IJ, Ballantyne JA. Does body mass index affect clinical outcome post-operatively and at five years after primary unilateral total hip replacement performed for osteoarthritis? A multivariate analysis of prospective data. *J Bone Joint Surg Br*. 2011;93:1178–82.
12. Dolan P. Modelling valuations for health states: the effect of duration. *Health Policy*. 1996;38:189–203.
13. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16:199–208.
14. Farrar JT, Portenoy RK, Berlin JA, Kinman JL, Strom BL. Defining the clinically important difference in pain outcome measures. *Pain*. 2000;88:287–294.
15. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001;94:149–158.
16. Franklin PD, Harrold L, Ayers DC. Incorporating patient-reported outcomes in total joint arthroplasty registries: challenges and opportunities. *Clin Orthop Relat Res*. 2013;471:3482–3488.

17. Franklin PD, Lewallen D, Bozic K, Hallstrom B, Jiranek W, Ayers DC. Implementation of patient-reported outcome measures in U.S. total joint replacement registries: Rationale, status, and plans. *J Bone Joint Surg Am.* 2014;96:104–109.
18. Gandhi R, Dhotar H, Davey JR, Mahomed NN. Predicting the longer-term outcomes of total hip replacement. *J Rheumatol.* 2010;37:2573–2577.
19. Gromov K, Greene ME, Sillesen NH, Troelsen A, Malchau H, Huddleston JI, Emerson R, Garcia-Cimbrello E, Gebuhr P. Regional differences between US and Europe in radiological osteoarthritis and self-assessed quality of life in patients undergoing total hip arthroplasty surgery. *J Arthroplasty.* 2014;29:2078–2083.
20. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *J Bone Joint Surg Am.* 1969;51:737–755.
21. Hofstede SN, Gademan MGJ, Vliet Vlieland TPM, Nelissen RGHH, Marang-van de Mheen PJ. Preoperative predictors for outcomes after total hip replacement in patients with osteoarthritis: a systematic review. *BMC Musculoskelet Disord.* 2016;17:212.
22. Holtzman J, Saleh K, Kane R. Effect of baseline functional status and pain on outcomes of total hip arthroplasty. *J Bone Joint Surg Am.* 2002;84-A:1942–1948.
23. Judge A, Javaid MK, Arden NK, Cushnaghan J, Reading I, Croft P, Dieppe PA, Cooper C. Clinical tool to identify patients who are most likely to achieve long-term improvement in physical function after total hip arthroplasty. *Arthritis Care Res (Hoboken).* 2012;64:881–889.
24. Keurentjes JC, Fiocco M, So-Osman C, Onstenk R, Koopman-Van Gemert AWMM, Pöhl RG, Kroon HM, Vliet Vlieland TPM, Nelissen RG. Patients with severe radiographic osteoarthritis have a better prognosis in physical functioning after hip and knee replacement: a cohort-study. *PLoS One.* 2013;8.
25. Kurtz S. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am.* 2007;89:780.
26. Leopold SS, Porcher R. Editorial: The minimum clinically important difference—the least we can do. *Clin Orthop Relat Res.* 2017;475:929–932.
27. Liu SS, Buvanendran A, Rathmell JP, Sawhney M, Bae JJ, Moric M, Perros S, Pope AJ, Poultsides L, Della Valle CJ, Shin NS, McCartney CJL, Ma Y, Shah M, Wood MJ, Manion SC, Sculco TP. A cross-sectional survey on prevalence and risk factors for persistent postsurgical pain 1 year after total hip and knee replacement. *Reg Anesth Pain Med.* 37:415–422.
28. Liu SS, Buvanendran A, Rathmell JP, Sawhney M, Bae JJ, Moric M, Perros S, Pope AJ, Poultsides L, Della Valle CJ, Shin NS, McCartney CJL, Ma Y, Shah M, Wood MJ, Manion SC, Sculco TP. Predictors for moderate to severe acute postoperative pain after total hip and knee replacement. *Int Orthop.* 2012;36:2261–2267.
29. MacWilliam CH, Yood MU, Verner JJ, McCarthy BD, Ward RE. Patient-related risk factors that predict poor outcome after total hip replacement. *Health Serv Res.* 1996;31:623–638.
30. Maltenfort M, Díaz-Ledezma C. Statistics In Brief: minimum clinically important difference—availability of reliable estimates. *Clin Orthop Relat Res.* 2017;475:933–946.
31. Mannion AF, Impellizzeri FM, Naal FD, Leunig M. Women demonstrate more pain and worse function before THA but comparable results 12 months after surgery. *Clin Orthop Relat Res.* 2015;473:3849–3857.
32. McGlothlin AE, Lewis RJ. Minimal clinically important difference: defining what really matters to patients. *JAMA.* 2014;312:1342–1343.
33. Naal FD, Impellizzeri FM, Leunig M. Which is the best activity rating scale for patients undergoing total joint arthroplasty? *Clin Orthop Relat Res.* 2009;467:958–965.
34. Nilsson AK, Aurell Y, Siösteen AK, Lohmander LS, Roos HP. Radiographic stage of osteoarthritis or sex of the patient does not predict one year outcome after total hip arthroplasty. *Ann Rheum Dis.* 2001;60:228–232.
35. Riediger W, Doering S, Krismer M. Depression and somatisation influence the outcome of total hip replacement. *Int Orthop.* 2010;34:13–18.
36. Rolfson O, Dahlberg LE, Nilsson J-A, Malchau H, Garellick G. Variables determining outcome in total hip replacement surgery. *J Bone Joint Surg Br.* 2009;91:157–161.
37. Shervin N, Dorrwachter JM, Bragdon CR, Malchau H. Comparison of Physician and Patient Administered Harris Hip Score. *J Arthroplasty.* 2009;24:e75.
38. Stambough JB, Xiong A, Baca GR, Wu N, Callaghan JJ, Clohisy JC. Preoperative joint space width predicts patient-reported outcomes after total hip arthroplasty in young patients. *J Arthroplasty.* 2016;31:429–433.
39. Tilbury C, Holtslag MJ, Tordoir RL, Leichtenberg CS, Verdegaal SHM, Kroon HM, Fiocco M, Nelissen RGHH, Vliet Vlieland TPM. Outcome of total hip arthroplasty, but not of total knee arthroplasty, is related to the preoperative radiographic severity of osteoarthritis. A prospective cohort study of 573 patients. *Acta Orthop.* 2016;87:67–71.
40. Valdes AM, Doherty SA, Zhang W, Muir KR, Maciewicz RA, Doherty M. Inverse relationship between preoperative radiographic severity and postoperative pain in patients with osteoarthritis who have undergone total joint arthroplasty. *Semin Arthritis Rheum.* 2012;41:568–575.
41. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care.* 1992;30:473–483.