Osteoarthritis and Cartilage 19 (2011) 155–162

Osteoarthritis and Cartilage

What is a good patient reported outcome after total hip replacement?

N.K. Arden †‡*, A. Kiran †, A. Judge †, L.C. Biant §‖, M.K. Javaid †, D.W. Murray †, A.J. Carr †, C. Cooper †‡, R.E. Field ‖¶

† Oxford NIHR Musculoskeletal Biomedical Research Unit, University of Oxford, Windmill Road, Headington, Oxford OX3 7LD, UK
‡ MRC Epidemiology Resource Centre, University of Southampton, Southampton General Hospital, Southampton SO16 6YD, UK
§ Royal Infirmary of Edinburgh, Little France Crescent, Old Dalkeith Road, Edinburgh EH16 4SA, UK
‖ Research & Education South West London Elective Orthopaedic Centre, Dorking Road, Epsom, Surrey KT18 7EG, UK
¶ St. Helier Hospital, Wrythe Lane, Carshalton, Surrey SM5 1AA, UK

ARTICLE INFO

Article history:
Received 7 April 2010
Accepted 4 October 2010

Keywords:
Hip replacement
Hip arthroplasty
Oxford Hip Score
Satisfaction
Patient-reported outcomes

SUMMARY

Objectives: There is an increasing movement to collect and report patient reported outcome measures (PROM's) following total hip replacement (THR). In the UK, the procedure specific PROM of choice is the Oxford Hip Score (OHS). It is currently unclear how to use this information to determine outcome following surgery. The aim of this study was to define a threshold for the OHS that is correlated with patient satisfaction.
Design: Prospective cohort study.
Setting: A district general hospital (St. Helier Hospital, Carshalton, UK).
Participants: 799 patients receiving THR from 1995 to 2004.
Main outcome measures: At 12 and 24 months after surgery patients were asked if they were satisfied with surgery and completed the OHS. Receiver operating characteristic (ROC) analyses were used to identify thresholds of follow-up OHS, which best discriminated patient satisfaction. Analyses were stratified by age, sex, body mass index (BMI), baseline OHS and patient expectations.
Results: 91.9% of patients were satisfied with THR at 12 months (92.8% at 24 months). Using the ROC technique, the OHS at 12 months associated with patient satisfaction was 38 and at 24 months 33. The OHS at 24 months associated with satisfaction was higher in those with highest tertile of baseline OHS (30, 33, 43 respectively), and lowest tertile of BMI.
Conclusions: We have identified a value of the OHS that predicts patient satisfaction 12–24 months following THR within a standard clinical setting. However, this threshold is markedly influenced by pre-operative OHS and should be stratified accordingly.

Introduction

Every year in the UK, over 70,000 total hip replacements (THR's) are performed[1], with 500,000 per year in the US[2], predominantly for osteoarthritis[3]. The number is set to increase over the next few years due to an increasing elderly population. The operation is thought to have an excellent outcome, with success rates of over 90% reported at 10 years[4]. However, success in these studies is defined by survival of the prosthesis, whereas there is an increasing movement to assess outcome after surgery using patient reported outcome measures (PROM's), which may yield lower success rates than the currently quoted figures using prosthesis survival.

The most commonly used outcome following THR in the UK is the Oxford Hip Score (OHS). It was introduced in 1996 as a PROM predominantly for use in clinical trials[5]. It is joint specific and has been extensively assessed for reliability and validity[6,7]. It is widely used and has been adopted as the outcome of choice for both the National Joint Registry (NJR) and the UK Government PROM initiative. The OHS was originally designed for use in clinical studies to measure population based changes rather than individual changes in outcome. One of the most important patient related outcomes is satisfaction with surgery, where an appreciable minority of patients have been shown to be unsatisfied with the outcome of their THR surgery[8].

The aim of this study was examine whether it is possible to define a post-operative OHS threshold that discriminated patient satisfaction with the operation.

* Address correspondence and reprint requests to: Nigel K. Arden, Oxford NIHR Musculoskeletal Biomedical Research Unit, University of Oxford, Windmill Road, Headington, Oxford OX3 7LD, UK. Tel: 44-(0)-2380-764027; Fax: 44-(0)-2380-777624.
E-mail address: nka@mrc.soton.ac.uk (N.K. Arden).

1063-4584/$ – see front matter © 2010 Published by Elsevier Ltd on behalf of Osteoarthritis Research Society International.
doi:10.1016/j.joca.2010.10.004

Methods

In 1999 an outcome programme was established for all patients undergoing THR at St. Helier Hospital, Carshalton. This is a busy district general hospital serving a local population of about 320,000 in the London boroughs of Sutton and Merton. The programme was designed to monitor the progress of patients whose operations were undertaken since the beginning of 1995. Whilst the study was formally active from 1995 to 2004, data has been collected outside of this range due to backtracking of patients into the database (pre-1995) and extraction of St. Helier patients from a large merged south west cohort (post 2004). The range of operation dates are 1986–2007, with median interquartile range (IQR) of 2002 (2000, 2004); 87% of operations occurred between 1995 and 2004.

All patients admitted for a THR completed a pre-operative questionnaire including an OHS and were subsequently sent annual postal questionnaires to assess outcome and satisfaction. Those patients that received an elective THR were recruited into the study; emergency THR patients were excluded. Subjects with previous hip fractures were also excluded. The review programme was run by members of the orthopaedic research and outcome unit and was independent of the patients continuing clinical care. The hip replacement surgery was performed by multiple consultant orthopaedic surgeons and their supervised trainees. The operating surgeon was not directly involved in the collection of data for the review programme. Pre-operative information was collected on patient age, gender, height and weight [from which body mass index (BMI) was calculated]. Patients were also asked the following questions regarding their expectations of surgery: (1) "*How painful do you expect your hip to be when you are fully recovered from this surgery? - Not at all painful, slightly painful, very painful*", (2) "*How limited do you expect to be in your usual activities, when you are fully recovered from this surgery? - Not limited at all, slightly limited, moderately limited, greatly limited*". They were also asked "*How long did it take before you recovered fully from your operation?*".

At the 12 and 24-month follow-up visits, in addition to the OHS, subjects were asked "*Are you satisfied with the result of your hip replacement*?" - the patients answered Yes or No to each question. The OHS consists of 12 questions asking patients to describe their hip pain and function during the past 4 weeks[5]. Each question is on a Likert scale taking values from 0 to 4. An overall score is created by summing the responses to each of the 12 questions. The total score can range from 0 to 48, where 0 is the worst possible score indicating severe hip symptoms and 48 is the best score suggesting excellent joint function. The *change* in OHS was the follow-up score minus the baseline score, and *percentage of potential improvement* (*PoPI*) defined as the percentage of the maximum potential benefit that the patient could achieve (follow-up score minus baseline score, divided by 48 minus baseline score, multiplied by 100).

Statistical methods

All statistical analysis was conducted using Stata SE v10, StatCorp, College Station, TX, USA and Matlab R2009b, The MathWorks, Natick, MA, USA.

Outcome variables

Analyses are restricted to patients receiving primary THR. If a patient had the procedure on both sides ($n = 95$ out of 799), we included the earliest operation, and bilateral operations were excluded. Outcomes of interest are the OHS at 12 and 24-month follow-up, and the difference between baseline and follow-up scores. However, these are continuous outcome variables, and instead we wish to create binary outcomes, by identifying the cut-off point on the OHS that relates to patient satisfaction. The clinical anchor for this analysis was whether or not the patient was satisfied with surgery. Hence, we aimed to identify the cut-point on the follow-up OHS, and change in OHS, that relates to patient satisfaction with surgery.

Two different statistical methods were used to identify cut-points for patient satisfaction. First, a receiver operating characteristic (ROC) curve analysis was used, where the gold standard is whether or not a patient improved according to the anchoring question, and we identify the cut-point on the follow-up OHS and change in OHS that maximises sensitivity and specificity. We validated the ROC results with a further statistical approach, the 75th percentile approach[11], which identifies the cut-off point corresponding to the 75th percentile of the follow-up OHS and change in OHS in patients answering yes to the anchoring question (satisfied with surgery).

Exposure/stratification variables

The methods described above to identify cut-points for the OHS, were then repeated stratifying by the following variables: age (tertiles), gender, BMI (tertiles), baseline OHS (tertiles), pre-operative expectations of pain (not at all painful vs any pain), and expectations of function (not limited at all vs any limitation) (see Table II).

Results

The cohort consisted of 799 patients at baseline, of whom 619 (77.5%) completed the 12-month follow-up questionnaire and 639 (80.0%) completed the 24-month questionnaire. Figure 1 shows the participants at baseline, 12 months and 24 months. Table I compares the clinical characteristics of all the patients at baseline with those that also attended the 24-month follow-up. We can see that the two groups were similar with no significant differences in the important patients' characteristics. Data on the underlying diagnosis was available on 239 patients only and revealed that 95.4% (228/239) had received their hip operation because of osteoarthritis/coxarthrosis. The remaining 4.6% (11/239) had THR due to avascular necrosis (1.67%), failed post fracture fixation (1.26%), acetabular erosion secondary to hemiarthroplasty (0.42%), unspecified arthritis (0.42%), hip dysplasia (0.42%) and joint pain (0.42%).

At 12 months 91.9% ($n = 564/614$) of patients were satisfied with their operation, with 92.8% ($n = 582/627$) at 24 months. Of the 45 patients who were not satisfied with their operation at 24 months,

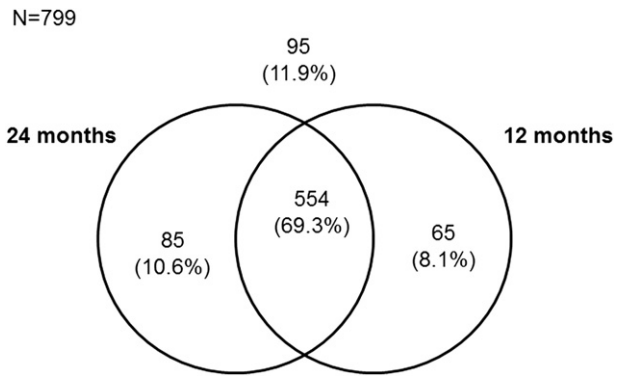

Fig. 1. Venn diagram showing the number of participants involved in the study (799 subjects recruited at baseline).

Table I
Baseline clinical characteristics of the patients

	Full cohort ($n=799$)	Non-respondents at 12-month follow-up ($n=180$)	Respondents at 12-month follow-up ($n=619$)	Non-respondents at 24-month follow-up ($n=160$)	Respondents at 24-month follow-up ($n=639$)
Age					
All, median (IQR)	68 (58,76) ($n=797$)	65 (54,74) ($n=179$)	68 (59,76) ($n=618$)*	67.5 (56,77) ($n=158$)	68 (59,75) ($n=639$)
Low tertile, median [range]	54 [20,62] ($n=269$)	51 [20,60] ($n=65$)	55 [27,63] ($n=213$)***	53 [20,60] ($n=54$)	55 [25,63] ($n=230$)**
Medium tertile, median [range]	68 [63,73] ($n=275$)	65 [63,71] ($n=56$)	69 [64,73] ($n=204$)***	68 [61,74] ($n=52$)	68 [64,73] ($n=216$)
High tertile, median [range]	79 [74,100] ($n=253$)	78.5 [72,100] ($n=58$)	79 [74,93] ($n=201$)	80 [75,100] ($n=52$)	79 [74,93] ($n=193$)
Gender					
Female, frequency (%)	480 (60.1%)	107 (59.4%)	373 (60.3%)	96 (60.0%)	384 (60.1%)
Male, frequency (%)	319 (39.9%)	73 (40.6%)	246 (39.7%)	64 (40.0%)	255 (39.9%)
BMI					
All, median (IQR)	27 (24,30) ($n=487$)	27 (24,30) ($n=106$)	27 (24,30) ($n=381$)	27 (23,30) ($n=93$)	27 (24,30) ($n=394$)
Low tertile, median [range]	23 [15,25] ($n=197$)	23 [15,25] ($n=42$)	23 [17,25] ($n=155$)	22 [15,24] ($n=32$)	24 [17,25] ($n=155$)**
Medium tertile, median [range]	27 [26,29] ($n=143$)	27 [26,29] ($n=33$)	27 [26,29] ($n=110$)	27 [25,29] ($n=34$)	27 [26,29] ($n=119$)
High tertile, median [range]	32 [30,43] ($n=147$)	33 [30,42] ($n=31$)	32 [30,43] ($n=116$)	33 [30,39] ($n=27$)	32 [30,43] ($n=120$)
OHS					
All, median (IQR)	17 (11,23) ($n=799$)	17 (12,24) ($n=180$)	17 (11,23) ($n=619$)	17 (12,23) ($n=160$)	17 (11,23) ($n=639$)
Low tertile, median [range]	9 [0,13] ($n=278$)	10 [1,15] ($n=71$)	9.5 [0,13] ($n=228$)*	10 [1,14] ($n=57$)	9 [0,13] ($n=230$)
Medium tertile, median [range]	17 [14,21] ($n=277$)	18 [16,21] ($n=51$)	17 [14, 21] ($n=205$)	17 [15,20] ($n=50$)	17 [14,21] ($n=212$)
High tertile, median [range]	27 [22,41] ($n=244$)	26.5 [24,30] ($n=58$)	27 [22,41] ($n=186$)	25 [21,36] ($n=53$)	27 [22,40] ($n=197$)*
Duration of pain, median (IQR)	1–3 years (1–3 years, 3–5 years) ($n=630$)	1–3 years (1–3 years, 3–5 years) ($n=83$)	1–3 years (1–3 years, 3–5 years) ($n=547$)	No observations	1–3 years (1–3 years, 3–5 years) ($n=630$)

*$p<0.05$, **$p<0.01$, ***$p<0.00005$.
Selection bias determined at each follow-up year by comparing non-respondents with respondents; for normally distributed data the independent 2 tail *t*-test was used; for non-normally distributed data the Wilcoxon rank sum test was used; chi-square test used for categorical data.

their respective median (IQR) follow-up OHS at 12 and 24 months was 27(17,35) and 23(17,29) respectively. 52.5% had reported a decrease in pain and 35.6% an increase in function (20% reported both). Table II presents the proportions of patients satisfied with surgery at each follow-up, broken down by baseline clinical characteristics. Using chi-square tests of association there was no evidence that baseline characteristics were associated with satisfaction scores. A chi square test for trend was also used to compare

Table II
Proportion of patients satisfied with surgery and their OHS, stratified by baseline clinical characteristics (tertile range as defined in Table I)

	Satisfaction		Baseline OHS	
	12 months Number (%)	24 months Number (%)	12 months Median (IQR)	24 months Median (IQR)
Age groups (tertiles)				
Low	195 (91.98%)	211 (92.95%)	18 (13,24), $n=213$	18 (13,24), $n=230$
Medium	189 (92.65%)	197 (94.26%)	16 (11,22.5), $n=204$	16.5 (10.5,23), $n=216$
High	179 (90.86%)	174 (91.10%)	15 (10,22), $n=201$	15 (10,22), $n=193$
p-value	0.805	0.472	0.0146*	0.0308*
Gender				
Female	340 (92.14%)	344 (91.98%)	15 (10,22), $n=373$	15.5 (10,22), $n=384$
Male	224 (91.43%)	238 (94.07%)	18 (12,24), $n=246$	19 (13,25), $n=255$
p-value	0.752	0.319	0.0003**	0.0001**
BMI (tertiles)				
Low	144 (93.51%)	142 (93.42%)	18 (12,24), $n=156$	18 (12,24), $n=155$
Medium	97 (88.18%)	105 (89.74%)	17 (11,25), $n=110$	17 (11,25), $n=119$
High	109 (93.97%)	109 (92.37%)	14 (10,20), $n=116$	15 (10,20), $n=120$
p-value	0.192	0.536	0.0017**	0.0127*
Baseline OHS (tertiles)				
Low	210 (92.51%)	209 (92.48%)	9.5 (7,12), $n=228$	9 (11,7), $n=230$
Medium	186 (91.63%)	194 (91.94%)	17 (16,20), $n=205$	17 (16,20), $n=212$
High	168 (91.30%)	179 (94.21%)	27 (24,30), $n=186$	27 (24,30), $n=197$
p-value	0.896	0.659	0.0001**	0.0001**
Pre-operative expectations - pain				
Not at all painful	384 (90.78%)	394 (92.71%)	17 (11,23), $n=428$	17 (11,23), $n=434$
Any pain	175 (94.09%)	184 (92.9%)	17 (11,23), $n=186$	18 (11,23), $n=201$
Chi square test, *p*-value	0.171	0.960	0.7001	0.8277
Pre-operative expectations - function				
Not limited at all	292 (93.29%)	284 (92.51%)	18 (12,24), $n=318$	18 (11,24), $n=314$
Any limitation	268 (90.24%)	294 (93.04%)	16 (10,22), $n=297$	16 (10,22), $n=321$
p-value	0.169	0.877	0.0132*	0.0028**

For non-normally distributed continuous data the Kruskal–Wallis rank sum test was used; the chi-square test used for categorical data. *$p<0.05$, **$p<0.01$, ***$p<0.00005$.

Table III
Outcomes at 12 and 24 months post-operation

	12 months ($n = 619$)	24 months ($n = 639$)
Baseline OHS, median (IQR)	17 (11,23)	17 (11,23)
Follow-up OHS, median (IQR)	43 (36,47)	44 (38,47)
Change in OHS from baseline, median (IQR)	24 (16,30)	24 (17,31)
PoPI, median (IQR)	83.3% (60.6%, 96.0%)	86.3% (65.6%, 97.1%)
Satisfied, (%)	91.9% (564/614)	92.8% (582/627)
Improved function, (%)	89.3% (550/616)	90.7% (572/631)
Reduced pain, (%)	94.3% (541/574)	92.7% (549/592)
Less medication, (%)	87.7% (515/587)	89.0% (544/611)
Improved since last visit, (%)	76.0% (562/608)	51.9% (329/634)

satisfaction of patients over time, by using satisfaction (yes/no) at year 1 as the binary outcome and tertiles of operation dates as the ordered categorical exposure. This was repeated using satisfaction responses at year 2. The results at each year of follow-up were non-significant (year 1, $p = 0.215$; year 2, $p = 0.488$), suggesting there was no linear trend present in the satisfaction responses of patients over time.

The changes in OHS from baseline to 12 months and 24 months and PoPI are shown in Table III. The median improvement in OHS was 24 at 12 months and 24 months. The percentage of patients reporting an improvement since their previous visit was 76.0% at 12 months, but this fell to 51.9% between 12 and 24 months.

The OHS at 12 months associated with patient satisfaction was ≥38 using the ROC curve technique (≥38 using the 75th centile approach). For the 24-month OHS the cut-point was 33 (ROC) and 40 (75th centile). Plots for both methods at 24 months can be seen in Fig. 2. The sensitivity and specificity of the ROC technique at 24 months are 89.7% and 86.7% respectively. The changes in OHS associated with satisfaction using the ROC were 15 at 12 months and 14 at 24 months.

Figure 3 displays the cut-points for the follow-up OHS, the change in OHS and PoPI, associated with satisfaction at 24 months using the ROC method, stratified by important pre-operative clinical variables. For the follow-up OHS (Fig. 3), there were little differences for age or gender at the 24-month visit. Patients with a high baseline OHS had higher follow-up OHS cut-points for satisfaction. Patients with the lowest BMI had the highest OHS threshold at 24 months (43, 32 and 34 respectively). Patients' pre-operative expectations had little impact on cut-points at 24 months.

The cut-points using change in OHS between baseline and 24 months (Fig. 3) demonstrated much greater variation between the clinical strata than those using the 24-month OHS. The value associated with satisfaction was higher in females than males and there were large variations according to BMI and the patients' pre-operative expectations. Patients with the highest baseline OHS required the lowest change in OHS to be satisfied. The PoPI have less variation than the changes score, but more than the score at 24

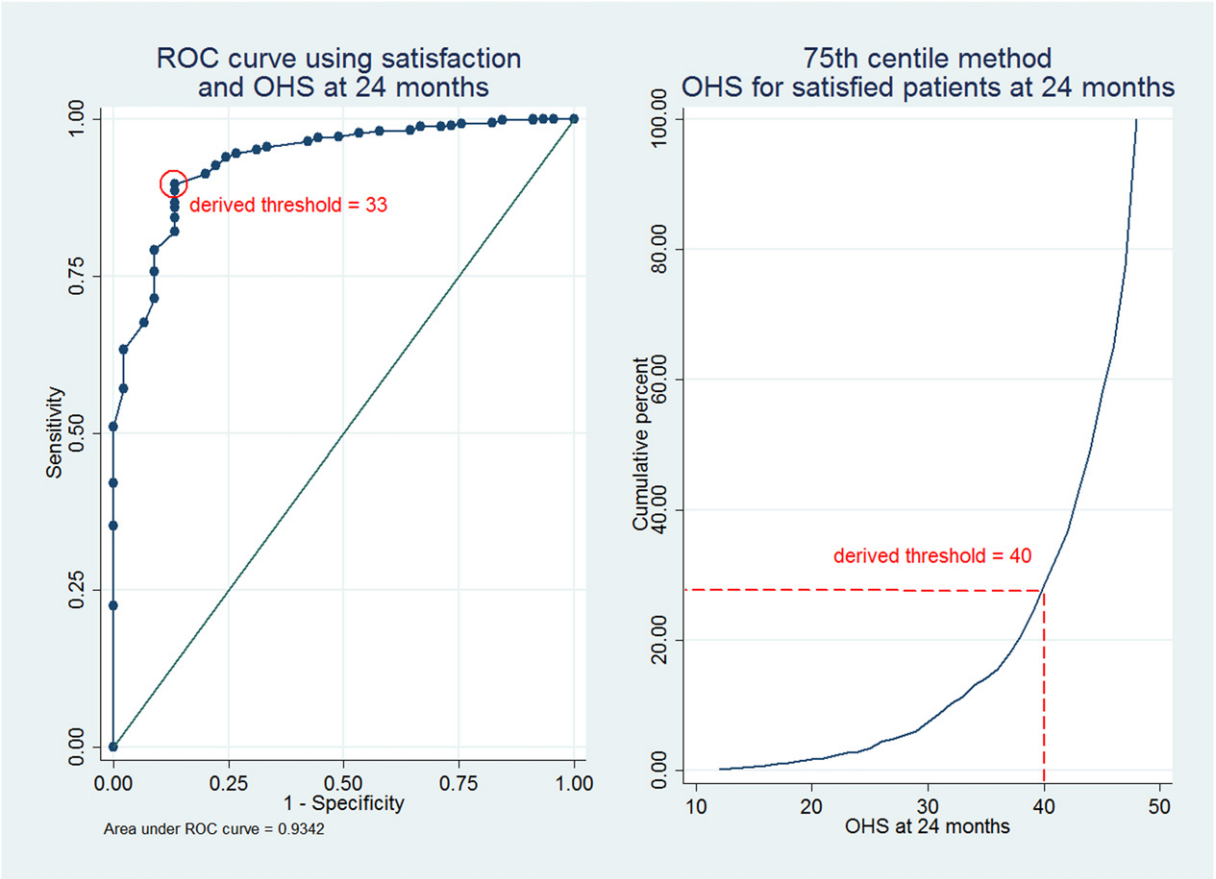

Fig. 2. ROC and 75th centile plots for satisfaction at 24 months. The 75th centile method cuts the ranked data such that positive responses are contained within 75% of the data. In our analysis, as '0' represents a poor result and '48' represents a good result, the 75th percentile has been taken from the top end of the OHS, i.e., the cut point is at 25%. The work of other authors may have their scoring system in reverse, where '48' is a poor result and '0' is a good result, in which case their cut point is at 75%.

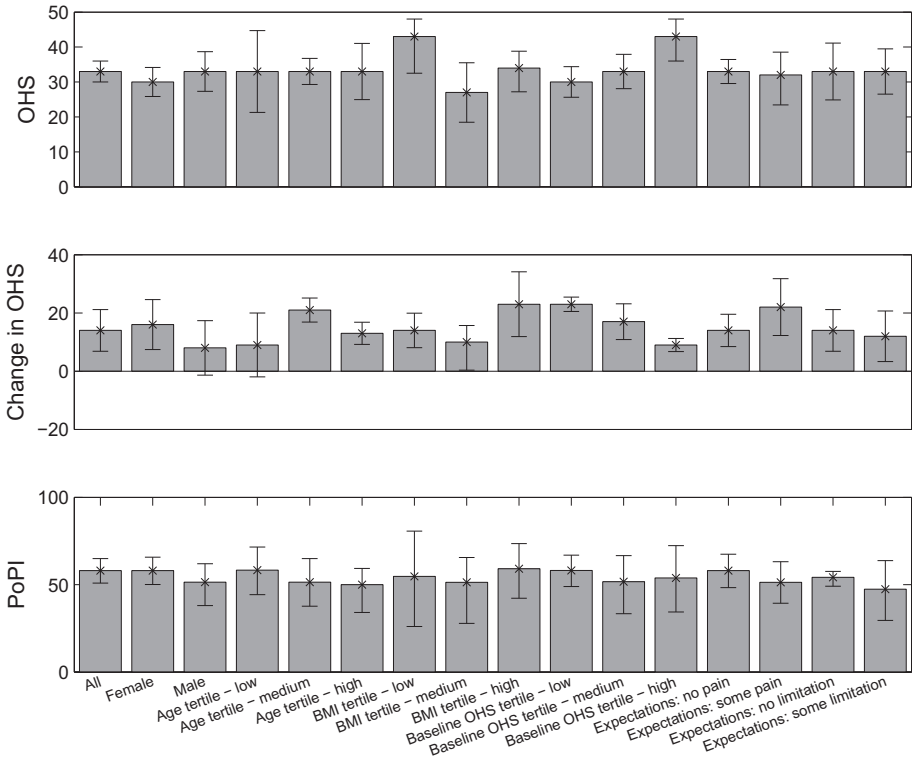

Fig. 3. ROC analysis: OHS based cut points using satisfaction as a binary classifier with 95% CI, stratified by baseline clinical variables.

months. Particularly, there were marked variations for gender and age.

Discussion

Principal findings

The OHS is currently being used as a PROM instrument, in order to assess whether hip replacement surgery has been successful from the patients perspective in terms of self reported pain and function. However the OHS is measured on a continuous scale from 0 (bad) to 48 (good), and it is unclear what cut point represents a satisfactory outcome when assessing patients' after surgery. For the first time, this study provides data on thresholds for the OHS that relate to patient satisfaction with surgery in standard practice in the NHS. Over 90% of patients are satisfied with their operation from 12 months, with the vast majority remaining satisfied at 24 months. Furthermore, even though patients reported continued improvement between 12 and 24 months, levels of satisfaction and change in OHS remained stable. We have identified a value of the OHS and change in OHS, which predicts patient satisfaction at 12 and 24 months within a standard clinical setting. The clinical application of this threshold is limited, however, as it varies according to several important factors such as pre-operative OHS and BMI and threshold stratified by these factors should be adopted.

Strength and limitations

The strengths of this study include the relatively large cohort, the use of a reliable, valid and responsive instrument for assessing outcomes of hip replacement surgery[6], data that has been collected prospectively with a good rate of follow-up within a standard NHS setting with multiple surgeons. While the benefits of this study were the comprehensive nature of its data collection and therefore the validity of the results, the generalisability of this study is limited as all subjects were recruited from one hospital. A repeat of this study at other centres would provide a better representation of the UK. As with most cohort studies, this study is subject to an incomplete follow-up of patients. Our response rates however were acceptable at between 77 and 80% and we found no evidence that non-responders differed from responders in important variables and hence are confident that this will not have significantly affected our results.

In order to estimate the OHS score for satisfaction at follow-up, we used a single anchoring question on patient satisfaction with the outcome of surgery. This is a similar concept to the widely used Patient Acceptable Symptom State (PASS)[9–11]. This is not a true PASS, as our anchoring question enquires about satisfaction with an operation that will include aspects of the patients current symptoms, but also their prior level of symptoms in addition to their response and expectations of surgery. Whilst a Likert scale response to satisfaction would have provided more detail in the analysis, the binary response used in the OHS questionnaire was sufficient (and required) for ROC curve analysis. It has also been suggested that rather than using one global anchor question, separate anchors should be used for each domain (pain and function) of the outcome measure. In addition other outcomes such as hospital facilities, infection rates and peri-operative mortality are not captured by the OHS and are important measures that inform patient choice.

What is already known?

In this study over 90% of patients reported they were satisfied with the outcome of THR surgery at each follow-up. A large

prospective study of THR operations carried out in five regions of the UK found that 89% of patients were satisfied with the operation at 1-year follow-up[8]. Data from the NJR found that 89.8% of patients were satisfied at least 1 year after the surgery[12]. Long term follow-up of the Trent regional arthroplasty register suggests satisfaction rates of 94.1% at 5 years[13] and 96.9% at 10 years[4]. The study by Nilsdotter *et al.* of patients receiving primary THR for OA in Sweden found satisfaction rates of 96% at 3.6 years[14], but as the authors highlight, even if the patient reports a bad outcome in pain and function they may still be satisfied with surgery, and satisfaction is a wide concept, not necessarily relevant in outcome after THR. This is why we have attempted to find a threshold for follow-up and change in OHS (a valid, responsive instrument to assess THR outcomes) in patients that are satisfied with surgery, rather than looking at satisfaction alone.

Other studies in the literature also suggest that there are an important minority of patients that have poor outcomes of THR in terms of change in pain and function. MacWilliam *et al.* found that approximately 16% of patients report no change or increased pain at 6 months, and 24% of patients report no change or decreased physical function at 6 months[15]. Nilsdotter *et al.* found between 9 and 25% of patients did not respond to treatment dependant on the way patients were classified as non-responders to surgery[14]. In line with other studies, their results also show that at follow-up, patients still had worse pain and function scores[16,17] than age–sex matched controls not receiving THR, suggesting that patients do not return to the same levels of pain and function as those in the general population. Quintana *et al.* demonstrated that at 6-month posts surgery, 70% of patients were classified as responders in terms of pain and function[16]. Whilst others have published values for the MCID based on WOMAC scores and VAS scales[11,16], to our knowledge, no other studies have provided cut-points for the OHS that relates to patient satisfaction, and such information has been requested in order to determine the real clinical or subjective meaning in changes in OHS[6].

What does this study add?

Increasing attention has been focused on the use of patient-reported outcome measures (PROMS) and assessments of satisfaction in evaluating THR surgery[18]. In the United Kingdom, there is now a national requirement to report PROMS on all patients undergoing THR's[19] using a condition-specific instrument OHS, a generic instrument (EQ5D) and general patient-specific information[20]. The idea behind these PROMS instruments is to measure whether surgery has worked by seeing if the patient "feels" better. By publishing such data the idea is to improve the quality of care received. These data will feed into the patient choice agenda as patients will be able to access such information when making a decision about surgery. The idea behind giving patients choice, is that it allows them to make an informed decision about their treatment with their general practitioner (GP), that best meets their needs. This feeds into the idea of using informed decision-making (IDM) in patient–GP communication. With shared decision making the onus is on the GP to ensure that patients have all the information they need, that they understand their treatment options, and then help them to make an informed decision that corresponds with their preferences[21].

However, the OHS is a continuous outcome measure, and to know that having hip surgery will change your score by '*x*' points is not informative to either patients or clinicians. Instead it is preferable to categorize patients into those who do or do not improve after surgery, so that as part of patient–clinician decision-making, a patient will know the likely improvement in pain and function they can expect as a percentage. Part of the PROMS agenda is to publish this data, and it is necessary to know what cut-point represents a clinically acceptable outcome of surgery, so we can estimate the proportion of patients that have responded to surgery in each hospital.

The results of this study provide a value of 38 on the 12-month OHS (33 at 24 months) that relates to patient satisfaction with surgery in standard practice in the NHS. This data will increase the usefulness of the OHS as a PROM instrument and make the data more meaningful to both patients and clinicians. We would however, caution against the use of a single set cut-point for the OHS as our analyses suggested that patients with lower baseline OHS (worse pre-operative pain and function) and those with higher BMI had lower values for OHS cut-points, suggesting that such patient groups are more likely to be satisfied with the outcomes of surgery with smaller improvements in pain and function. We should therefore recommend using different set points based on specific pre-operative patient characteristics. These data provide the first published information on how to use the OHS for assessing patient satisfaction post surgery for use in clinical settings, however, validation in other cohorts is ideally required before their widespread implementation. The alternative approach of using change in OHS as an outcome, did not perform well in these analyses as it leads to more heterogeneity amongst strata.

Unanswered questions and future research

THR surgery is one of the best surgical interventions making a substantial contribution to public health. It is one of the most common elective surgical procedures, shown to be a cost effective, with good prosthesis survival rates, reducing pain, increasing mobility and improving quality of life[22–25]. Increasing attention is focused on the use of PROMS and assessments of satisfaction in evaluating THR surgery, with the OHS being the condition specific PROM of choice. This study has identified a threshold for follow-up OHS at 12 and 24 months that relates to a clinically important improvement related to patient satisfaction with surgery, with good sensitivity and specificity. Further research is required to validate these cut points for specific patient sub groups in other cohorts and to define the cut-points at 6-month post-operation.

Author contributions

MKJ, NKA, REF and AK were involved in the analysis and interpretation of data, and NKA, REF and MKJ in conception and design of the study. All authors were involved in drafting the manuscript. All authors had full access to all of the data in the study, and reviewed the manuscript. REF is the guarantor.

Role of the funding source

No sponsors participated in the design and conduct of the study.

Conflicts of interest

There are no conflicts of interest.

Acknowledgements

This study has received financial support from the Oxford Musculoskeletal NIHR BRU, University of Oxford.

Appendix

Results from the ROC analysis at 24 months; sensitivity, specificity, area under the curve (AUC) with 95% confidence interval (CI) and corresponding OHS related cut-point

	n	Sensitivity	Specificity	AUC	Cut point
OHS at 24 months					
All	627	0.897 (0.869, 0.920)	0.867 (0.732, 0.949)	0.934 (0.902, 0.967)	33
Female	374	0.933 (0.901, 0.957)	0.800 (0.614, 0.923)	0.919 (0.869, 0.968)	30
Male	253	0.916 (0.873, 0.948)	0.933 (0.681, 0.998)	0.959 (0.931, 0.987)	33
Age tertile - low	227	0.938 (0.897, 0.967)	0.813 (0.544, 0.960)	0.937 (0.888, 0.986)	33
Age tertile - medium	209	0.898 (0.848, 0.937)	0.917 (0.615, 0.998)	0.931 (0.859, 1.000)	33
Age tertile - high	191	0.845 (0.782, 0.895)	0.882 (0.636, 0.985)	0.936 (0.887, 0.984)	33
BMI tertile - low	152	0.732 (0.652, 0.803)	1.000 (0.692, 1.000)	0.895 (0.822, 0.967)	43
BMI tertile - medium	117	0.981 (0.933, 0.998)	0.917 (0.615, 0.998)	0.980 (0.949, 1.000)	27
BMI tertile - high	118	0.908 (0.838, 0.955)	1.000 (0.664, 1.000)	0.966 (0.934, 0.997)	34
Baseline OHS tertile - low	226	0.871 (0.818, 0.913)	0.941 (0.713, 0.999)	0.947 (0.915, 0.980)	30
Baseline OHS tertile - medium	211	0.933 (0.888, 0.964)	0.941 (0.713, 0.999)	0.975 (0.941, 1.000)	33
Baseline OHS tertile - high	190	0.726 (0.655, 0.790)	0.909 (0.587, 0.998)	0.898 (0.819, 0.977)	43
Expectations: no pain	425	0.909 (0.876, 0.935)	0.871 (0.702, 0.964)	0.942 (0.906, 0.979)	33
Expectations: some pain	189	0.880 (0.822, 0.924)	0.857 (0.572, 0.982)	0.909 (0.838, 0.981)	32
Expectations: no limitation	307	0.940 (0.906, 0.965)	0.826 (0.612, 0.950)	0.940 (0.901, 0.980)	33
Expectations: some limitation	307	0.849 (0.802, 0.889)	0.909 (0.708, 0.989)	0.936 (0.883, 0.989)	33
Change in OHS					
All	627	0.881 (0.852, 0.907)	0.778 (0.629, 0.888)	0.920 (0.884, 0.955)	14
Female	374	0.834 (0.791, 0.872)	0.800 (0.614, 0.923)	0.909 (0.863, 0.954)	16
Male	253	0.962 (0.929, 0.983)	0.800 (0.519, 0.957)	0.953 (0.911, 0.995)	8
Age tertile - low	227	0.967 (0.933, 0.987)	0.688 (0.413, 0.890)	0.904 (0.829, 0.979)	9
Age tertile - medium	209	0.670 (0.600, 0.735)	1.000 (0.735, 1.000)	0.886 (0.823, 0.949)	21
Age tertile - high	191	0.908 (0.855, 0.947)	0.941 (0.713, 0.999)	0.956 (0.913, 0.999)	13
BMI tertile - low	152	0.880 (0.815, 0.929)	0.900 (0.555, 0.997)	0.926 (0.851, 1.000)	14
BMI tertile - medium	117	0.962 (0.905, 0.990)	0.833 (0.516, 0.979)	0.936 (0.868, 1.000)	10
BMI tertile - high	118	0.706 (0.612, 0.790)	1.000 (0.664, 1.000)	0.899 (0.812, 0.985)	23
Baseline OHS tertile - low	226	0.842 (0.785, 0.889)	1.000 (0.805, 1.000)	0.946 (0.912, 0.979)	23
Baseline OHS tertile - medium	211	0.923 (0.876, 0.956)	0.941 (0.713, 0.999)	0.974 (0.944, 1.000)	17
Baseline OHS tertile - high	190	0.888 (0.833, 0.930)	0.909 (0.587, 0.998)	0.913 (0.816, 1.000)	9
Expectations: no pain	425	0.888 (0.853, 0.918)	0.839 (0.663, 0.945)	0.934 (0.894, 0.974)	14
Expectations: some pain	189	0.646 (0.570, 0.716)	1.000 (0.768, 1.000)	0.894 (0.823, 0.964)	22
Expectations: no limitation	307	0.887 (0.845, 0.922)	0.783 (0.563, 0.925)	0.916 (0.865, 0.967)	14
Expectations: some limitation	307	0.916 (0.877, 0.945)	0.773 (0.546, 0.922)	0.926 (0.877, 0.974)	12
PoPI					
All	627	0.864 (0.834, 0.891)	0.956 (0.849, 0.995)	0.954 (0.929, 0.978)	58%
Female	374	0.860 (0.819, 0.895)	0.933 (0.779, 0.992)	0.940 (0.904, 0.976)	58%
Male	253	0.912 (0.868, 0.945)	1.000 (0.782, 1.000)	0.978 (0.960, 0.996)	51%
Age tertile - low	227	0.910 (0.863, 0.945)	0.938 (0.698, 0.998)	0.958 (0.920, 0.995)	58%
Age tertile - medium	209	0.893 (0.842, 0.933)	0.917 (0.615, 0.998)	0.933 (0.874, 0.991)	51%
Age tertile - high	191	0.874 (0.815, 0.919)	1.000 (0.805, 1.000)	0.970 (0.947, 0.994)	50%
BMI tertile - low	152	0.887 (0.823, 0.934)	0.900 (0.555, 0.997)	0.933 (0.875, 0.991)	55%
BMI tertile - medium	117	0.914 (0.844, 0.960)	1.000 (0.735, 1.000)	0.979 (0.955, 1.000)	51%
BMI tertile - high	118	0.853 (0.773, 0.914)	1.000 (0.664, 1.000)	0.960 (0.920, 1.000)	59%
Baseline OHS tertile - low	226	0.833 (0.775, 0.880)	1.000 (0.805, 1.000)	0.948 (0.915, 0.980)	58%
Baseline OHS tertile - medium	211	0.938 (0.894, 0.968)	0.941 (0.713, 0.999)	0.975 (0.943, 1.000)	52%
Baseline OHS tertile - high	190	0.883 (0.826, 0.926)	0.909 (0.587, 0.998)	0.927 (0.856, 0.999)	54%
Expectations: no pain	425	0.881 (0.845, 0.911)	0.968 (0.833, 0.999)	0.966 (0.941, 0.991)	58%
Expectations: some pain	189	0.857 (0.796, 0.905)	0.929 (0.661, 0.998)	0.923 (0.865, 0.982)	51%
Expectations: no limitation	307	0.919 (0.881, 0.948)	0.957 (0.781, 0.999)	0.963 (0.934, 0.991)	54%
Expectations: some limitation	307	0.888 (0.845, 0.922)	0.909 (0.708, 0.989)	0.946 (0.904, 0.988)	47%

References

1. National Joint Registry. National Joint Registry for England and Wales. 5th Annual Report 2008.
2. United States Bone and Joint Decade: The Burden of Musculoskeletal Diseases in the United States. Rosemont, IL: American Academy of Orthopaedic Surgeons; 2008.
3. Culliford DJ, Maskell J, Beard DJ, Murray DW, Price AJ, Arden NK. Temporal trends in hip and knee replacement in the United Kingdom: 1991 to 2006. J Bone Joint Surg Br 2010;92(1):130–5.
4. Allami MK, Fender D, Khaw FM, Sandher DR, Esler C, Harper WM, *et al*. Outcome of Charnley total hip replacement across a single health region in England. The results at ten years from a regional arthroplasty register. J Bone Joint Surg Br 2006;88(10):1293–8.
5. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996;78-B(2):185–90.
6. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, *et al*. The use of the Oxford hip and knee scores. J Bone Joint Surg Br 2007;89-B(8):1010–4.
7. Field RE, Cronin MD, Singh PJ. The Oxford hip scores for primary and revision hip replacement. J Bone Joint Surg Br 2005;87(5):618–22.
8. Williams O, Fitzpatrick R, Hajat S, Reeves BC, Stimpson A, Morris RW, *et al*. Mortality, morbidity, and 1-year outcomes of

primary elective total hip arthroplasty. J Arthroplasty 2002;17 (2):165–71.
9. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? Ann Rheum Dis 2007;66(Suppl 1).
10. 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(2):102–9.
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 (1):29–33.
12. National Joint Registry. National Joint Registry for England and Wales. 2nd Annual Report 2005.
13. Fender D, Harper WM, Gregg PJ. Outcome of Charnley total hip replacement across a single health region in England: the results at five years from a regional hip register. J Bone Joint Surg Br 1999;81-B(4):577–81.
14. Nilsdotter AK, Petersson IF, Roos EM, Lohmander LS. Predictors of patient relevant outcome after total hip replacement for osteoarthritis: a prospective study. Ann Rheum Dis 2003;62 (10):923–30.
15. 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(5):623–38.
16. 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(12):1076–83.
17. Busija L, Osborne RH, Nilsdotter A, Buchbinder R, Roos EM. Magnitude and meaningfulness of change in SF-36 scores in four types of orthopedic surgery. Health Qual Life Outcomes 2008;6:55.
18. Professor the Lord Darzi of Denham KBE, HonFREng, FmedSci. High Quality Care for All. NHS Next Stage Review Final Report - Summary. Department of Health; 30-6-2008
19. Standard NHS Contract for Acute Services. Department of Health; 13-12-2007.
20. Guidance on the Routine Collection of Patient Reported Outcome Measures (PROMs). Department of Health; 2008.
21. Stacey D, Hawker G, Dervin G, Tomek I, Cochran N, Tugwell P, *et al*. Management of chronic pain: improving shared decision making in osteoarthritis. BMJ 2008;336(7650):954–5.
22. NHS Centre for Reviews and Dissemination. Total Hip Replacement. York: NHS Centre for Reviews and Dissemination; 1996. Effective Health Care Bulletin.
23. Chang RW, Pellisier JM, Hazen GB. A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. JAMA 1996;275(11):858–65.
24. Kirwan JR, Currey HL, Freeman MA, Snow S, Young PJ. Overall long-term impact of total hip and knee joint replacement surgery on patients with osteoarthritis and rheumatoid arthritis. Br J Rheumatol 1994;33(4):357–60.
25. Liang MH, Cullen KE, Larson MG, Thompson MS, Schwartz JA, Fossel AH, *et al*. Cost-effectiveness of total joint arthroplasty in osteoarthritis. Arthritis Rheum 1986;29 (8):937–43.