



Can patient-reported outcomes predict re-operations after total hip replacement?

Ted Eneqvist^{1,2,3}  · Szilárd Nemes^{1,2} · Erik Bülow^{1,2} · Maziar Mohaddes^{1,2,3} · Ola Rolfson^{1,2,3}

Received: 27 November 2017 / Accepted: 28 November 2017 / Published online: 3 January 2018
© The Author(s) 2018. This article is an open access publication

Abstract

Purpose We investigated if patient-reported outcomes (PROMs) one year after total hip replacement (THR) can predict the risk of re-operation using data from the Swedish Hip Arthroplasty Register.

Methods A total of 75,899 patients with THR due to osteoarthritis operated in 2002–2014 were included. We used Kaplan-Meier and Cox regression to investigate the relationship between one-year post-operative PROMs and risk of re-operation (all types of further hip surgery). The predictive power of the model and post-operative PROMs were evaluated by concordance index (C).

Results Kaplan-Meier estimates for not being re-operated at eight years was 95.5% (95%CI; 95.3–95.8). Cox regression analyses showed that all PROMs, except for EQ-VAS, were associated with re-operation. The full model had a concordance index of 0.68. Satisfaction (C = 0.65) and pain (C = 0.65) in isolation had the highest predictive power.

Conclusions Worse PROMs predicted higher risk of re-operation. Therefore, we believe PROMs may be helpful in identifying patients at risk for re-operation and timely address their problems.

Keywords Total hip replacement · Re-operation · Patient-reported outcome measurements · Prediction

Introduction

Total hip replacement (THR) is a well-established, cost effective method for patients with hip osteoarthritis when non-surgical treatment is ineffective. Survivorship of some prosthetics has been reported above 95% at 10 years [1, 2]. The Swedish Hip Arthroplasty register (SHAR) collects information from all of the primary THR procedures on demographic, surgical, and patient-reported outcome measures data, pre- and post-operatively [3].

An important minority operated with THRs will, unfortunately, experience complications, and some will need re-operation due to early or late complications. The most common reasons for re-operations are: infection, fracture, loosening or dislocation [3]. Such an event might be terrible for the individual patient, often resulting in impaired function and disability [4–6]. But these events are also very costly to society with increased medical costs, sick leave, etc. It has also been reported that results following revision surgery are less likely to be as good as the first operation [7].

Complications being uncommon after THR, few patients are in need of re-operation following a standard THR. Hence, the need for follow-ups after THR for all patients is being questioned [8, 9]. As a consequence, many healthcare providers have abandoned routine follow-ups after standard THRs.

There are several reports presenting risk factors for revision following THR, including patient-related factors [10–13], and implant and surgical procedure-related factors [12, 13], as well as PROMs-related factors such as the Harris Hip [14] and Oxford Hip scores [15, 16]. To our knowledge there are no reports using time-dependent area under the curve values (AUC-values)/concordance indices (C). Similarly, combinations of several PROMs, to investigate their association with

✉ Ted Eneqvist
ted.eneqvist@vgregion.se

¹ Swedish Hip Arthroplasty Register, Sögladsgatan 11, 41680 Gothenburg, Sweden

² Department of Orthopaedics, Institute of Clinical Sciences, the Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

³ Department of Orthopaedics, Sahlgrenska University Hospital, Gothenburg, Sweden

the prediction of the risk of a late re-operation following THR have not been used previously.

The aim of this study was to examine if patient-reported outcome measures (PROMs) reported at one-year after THR can predict the risk of future re-operation.

Method

In 1979, SHAR was started with the purpose to collect prospective observational data on all THRs performed in Sweden. The aim of the register was not only to gather data but also to compare results between clinics and longitudinally monitor the outcomes of the procedures with a focus on implants used, surgical techniques, and complications. The register covers both publicly and privately funded hospitals, and today the completeness of registrations has been reported at 98.3% [3]. In 2002, SHAR launched a nationwide PROMs program for elective THR patients with the purpose of complementing the traditional outcome variables with patient-reported outcomes for pain, function, and health-related quality of life [17, 18]. In 2008, all orthopaedic clinics in Sweden performing THRs joined the PROMs program. The PROMs program invites all patients scheduled for elective THR in Sweden to participate. Patients are asked to complete a short questionnaire pre-operatively, and a follow-up survey is mailed to patients at one, six, and ten years post-operatively. The response frequency has been reported at roughly 90%, both pre-operatively and at the one-year follow-up [3, 18].

Outcome measures

The outcome in this study is re-operation later than one year after the index surgical procedure. This includes for all reasons and all types of surgical procedures. The outcome (re-operation) was measured from one year following the index surgery until the occurrence of re-operation, death, or December 31st 2015, whichever came first. The predictors for re-operation are PROMs collected from the PROMs program in SHAR [19]. The PROMs program in SHAR includes the EQ-5D health status questionnaire [20], a hip pain visual analogue scale (VAS) [21], and at follow-ups, satisfaction with the outcome using a VAS. The EQ-5D questionnaire is internationally one of the most widely used generic health status measures [22]. The EQ-5D descriptive system includes five dimensions of health: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. In addition, the EQ-5D contains a VAS addressing general health (EQ VAS) where 0 represents the worst and 100 represents the best possible health state. The hip pain VAS addresses the patient's hip pain from a scale of 0 to 100, where 0 represents no pain and 100 the worst possible pain. At the follow-ups, satisfaction VAS addresses satisfaction with the outcome following THR. The

scale ranges from 0 to 100, where 0 represents very satisfied and 100 very dissatisfied. Included in the PROMs program are a patient-reported Charnley classification used to determine patient-reported musculoskeletal comorbidity. Similar to the original Charnley classification this patient-reported comorbidity question is classified as: class A corresponding to a unilateral hip disorder, class B being a bilateral hip disorder, and class C corresponding to a walking impairment due to multiple joint involvement or other medical comorbidities. To identify patients at the follow-ups the register uses the unique personal identity number (PIN) given to all inhabitants in Sweden.

Patient selection

We obtained data from the Swedish Hip Arthroplasty Register on patients operated with THR between 2002 and 2014. We excluded all primary procedures to the second hip in patients with bilateral THRs. We have also excluded all patients without lateral or posterior approaches, resurfacing prostheses, patients diseased or re-operated within one year, and patients with a non-osteoarthritis diagnosis (Fig. 1).

Statistical calculations

We investigated the demographics, surgical techniques, and post-operative PROMs between the two groups. We used Kaplan-Meier curves to summarize re-operation probabilities. Data was summarized as frequencies for categorical data, and means and standard deviations (SD) for continuous data. Comparisons of the re-operated and non-re-operated groups were conducted with the chi-square test for categorical, and the t-test for continuous, variables. We used a Cox regression analysis adjusted for post-operative PROMs, age and sex to investigate the association between post-operative PROMs, Charnley class, and age and sex with re-operation up to eight years following THR, after this there were too few re-operations to perform accurate analyses. The predictive power of PROMs one year post-operatively was evaluated by the concordance index (C) [23]. This value ranges from 0.5 to 1, with values closer to 1 being a better score. A value of 0.5 means that PROMs cannot predict the risk of re-operation, while a value of 1 indicates that re-operation could be completely predicted by the model.

Ethical considerations

Ethical approval was obtained from the Regional Ethical Review Board in Gothenburg, Sweden, registration number 268-17.

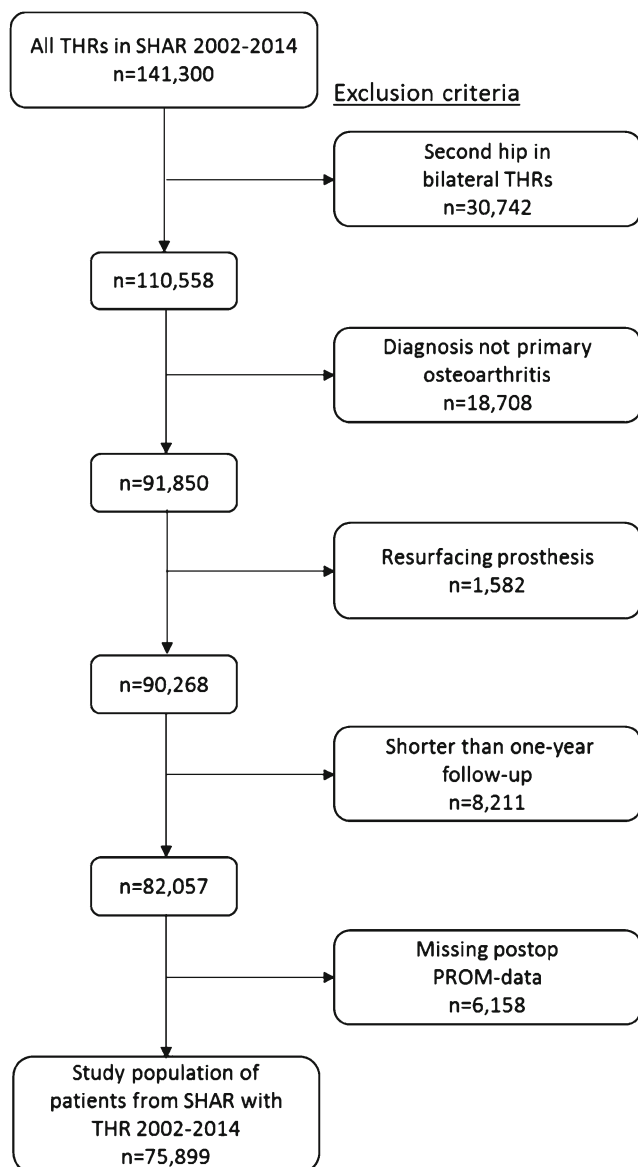


Fig. 1 Flowchart describing selection process. SHAR Swedish Hip Arthroplasty Register, THR total hip replacement, PROM patient-reported outcome measures

Results

Between 2002 and 2014 there were 141,300 patients operated with a THR. A total of 75,899 were included according to the selection criteria. 1,405 patients had a revision performed during this time period. The mean time until re-operation was 1,589 days (SD 887 days). At the ten-year follow-up the prosthesis survival rate was 95.5% (95.3–95.8) (Fig. 2). Except for the type of prosthesis, there were significant differences between the non- and re-operated groups (Table 1). Cox regression modeling showed an association with PROMs for risk of re-operation with pain ($B = 1.01$, 95% CI: 1.01, 1.02), satisfaction VAS ($B = 1.02$, 95% CI: 1.01, 1.02), and the EQ-5D-index ($B = 0.64$, 95% CI: 0.49, 0.85), but none for the EQ

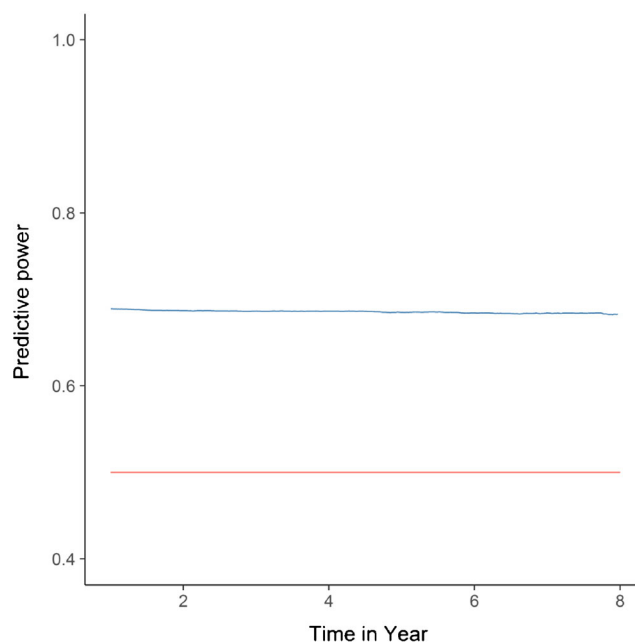


Fig. 2 Predictive power of the model. The blue line shows the change in predictive power of the model during the follow-up time. The red line indicates a predictive power of 0.5

VAS ($B = 1.00$, 95% CI: 0.99, 1.01). The Cox regression also presented an association for the risk of re-operation with age ($B = 0.98$, 95% CI: 0.97, 0.98), sex (female $B = 0.71$, 95% CI: 0.64, 0.79), Charnley class B ($B = 0.75$, 95% CI: 0.62, 0.91), and C ($B = 0.83$, 95% CI: 0.73, 0.94). The model had moderate predictive power, with a concordance index of 0.68 (Fig. 2). Satisfaction-VAS ($C = 0.650$) and pain VAS ($C = 0.649$) had the highest predictive powers: the EQ-5D ($C = 0.607$) and EQ VAS ($C = 0.602$) had a lower predictive capacity (Table 2).

Figure 3 shows an example of the probability of being re-operated for three different levels of PROMs (good, average and poor) following THR. The predictive power was relatively constant up to eight years post-surgery (Fig. 2). The observed re-operations and the predicted re-operation using the model above differed only marginally.

Discussion

Summary of findings

By using data from SHAR between 2002 and 2014 we have shown an association between one-year post-operative PROMs and the risk of future re-operation in patients operated with THR. There were significant differences in demographics, surgical data and post-operative PROMs between the group of patients that received a re-operation or not. The model we constructed in order to predict a risk of future re-

Table 1 Demographic data, surgical data and post-operative PROMs between the non-re-operated and re-operated groups

Demographic	Not re-operated	Re-operated	<i>p</i> -value
Cases, n	74,494	1405	
Age, mean (SD)	68.82 (9.9)	66.19 (10.0)	<0.001
Female, n (%)	42,359 (56.9)	712 (50.7)	<0.001
Prosthesis, n (%)			0.145
Cemented	53,092 (71.3)	1013 (72.1)	
Uncemented	11,274 (15.1)	205 (14.6)	
Hybrid	1283 (1.7)	34 (2.4)	
Reversed hybrid	8845 (11.9)	153 (10.9)	
Charnely class n (%)			<0.001
A	34,341 (46.1)	558 (39.7)	
B	8281 (11.1)	138 (9.8)	
C	31,872 (42.8)	709 (50.5)	
Pain VAS, mean (SD)	13.67 (17.7)	26.28 (25.5)	<0.001
Satisfaction VAS, mean (SD)	15.44 (20.0)	30.40 (30.6)	<0.001
EQ VAS, mean (SD)	76.29 (20.1)	68.55 (23.8)	<0.001
EQ-5D index, mean (SD)	0.79 (0.2)	0.67 (0.3)	<0.001
Mobility, n (%)			<0.001
1	44,922 (60.3)	618 (44.0)	
2	29,469 (39.6)	785 (55.9)	
3	101 (0.1)	2 (0.1)	
ADL, n (%)			<0.001
1	68,740 (92.3)	1200 (85.4)	
2	5312 (7.1)	197 (14.0)	
3	441 (0.6)	8 (0.6)	
Usual activities, n (%)			<0.001
1	57,682 (77.4)	883 (62.8)	
2	15,270 (20.5)	442 (31.5)	
3	1542 (2.1)	80 (5.7)	
Pain/discomfort, n (%)			<0.001
1	32,601 (43.8)	386 (27.5)	
2	38,442 (51.6)	850 (60.5)	
3	3450 (4.6)	169 (12.0)	
Anxiety/depression, n (%)			<0.001
1	58,295 (78.3)	931 (66.3)	
2	15,097 (20.3)	416 (29.6)	
3	1100 (1.5)	58 (4.1)	
Time to re-operation (days), mean (SD)	n/a	1589 (887)	

Categorical data presented as n (%) and continuous data as mean (standard deviation)

operation using post-operative PROMs had a moderate power to predict the risk of future re-operation: and there were small differences between the predicted and observed re-operation frequencies. For PROMs, satisfaction- and pain VAS had the highest power, and EQ-VAS the lowest power to predict the risk of future re-operation. Since our aim was to predict re-operation with as high accuracy as possible, we chose to keep EQ-VAS in the model favouring prediction over explanation [24].

Relation to other research

Earlier reports have shown that the strongest correlation with the risk of revision following THR is patient-related factors such as age and sex. In a systematic review by Prokopetz et al. [13] including 65 articles investigated with revision following THR for any reason as outcome, age (40%) and sex (28%) were the most common variables associated with predicting risk of future revision. This is also supported by later reports

Table 2 Association of post-operative PROMs, Charnley class, age and sex with re-operation using Cox regression adjusted for age, sex and postoperative PROMs. Predictive power of post-operative PROMs, Charnley class, age and sex with re-operation

Variables	HR	95% CI lower	95% CI upper	C
Charnley class A	Ref			0.539
Charnley class B	0.752	0.623	0.909	
Charnley class C	0.829	0.728	0.943	
Pain VAS	1.014	1.011	1.017	0.649
EQ VAS	1.001	0.998	1.005	0.602
Satisfaction VAS	1.016	1.013	1.018	0.650
EQ-5D index	0.643	0.489	0.846	0.607
Age	0.975	0.970	0.980	0.566
Male	Ref			0.535
Female	0.713	0.641	0.793	

HR hazard ratio, CI confidence interval, C concordance index

for sex [11, 13] and age [10–13]. In our study we also present an association with a higher risk of re-operation for younger patients and males, but these factors are weaker than post-operative PROMs to predict the risk of a late re-operation.

There are few other reports investigating the association of post-operative PROMs and the risk of future re-operation. Our study is consistent with those that have shown an association with post-operative PROMs and re-operation. Sing et al. have

shown that low Harris Hip scores at two- and five years are associated with a higher risk of revision at each time point [14]. Rothwell et al. have used data from the New Zealand Joint Registry and investigated the association between early revision and Oxford Hip scores (OHS) in 17,690 THRs. They found that for every one unit decrease in OHS at six months the risk for revision within two years increased by 9.7%. They also reported that 70% of the revisions would be captured if the patients that scored in the lowest 28% of the OHS had had extra follow-ups [15]. Devane et al. expanded the investigation by Rothwell et al. and suggested that patients who score poor OHS at six months and five years needed further follow-up in order to capture those in need of a revision in a timely manner [16]. In comparison to these studies we have a substantially larger number of patients in our calculation: and we also take several PROMs into consideration when performing our analysis. We have also selected our patients in order to reduce confounding biases and provide a homogenous group to investigate. The previously described studies presenting patient-related factors [10–13] and PROMs [14–16] to predict the risk of future revision following THR used OR (odds ratio) and HR (hazard ratio) to predict the risk of revision. But these methods have been declared unsuitable for such conclusions since they lack predictive strength [25]. Together with a Cox regression we have used the concordance index, which gives a substantially higher power to the predictive power of our analyses.

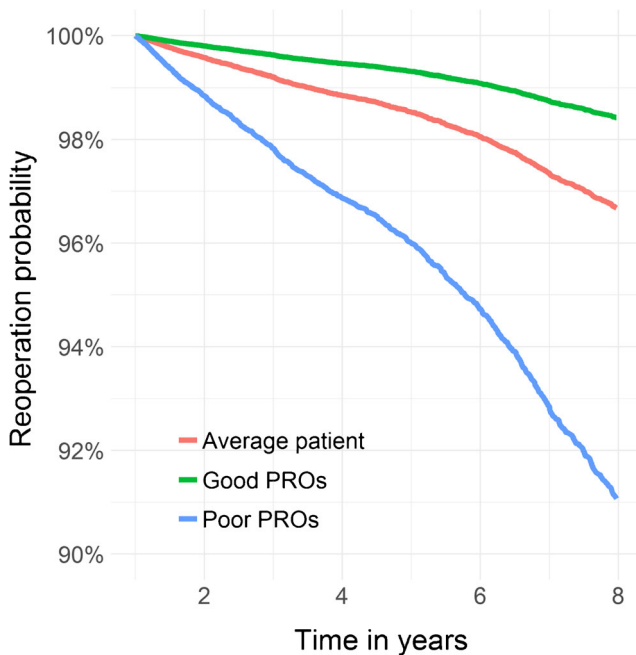


Fig. 3 Probability of being re-operated. Cox regression survival estimates for the average patient (average baseline characteristics and 1 year PROMs), a patient with good PROs (no hip VAS pain, full satisfaction, EQ mobility = 1 EQ pain/discomfort = 1) and a patient with poor PROs (hip VAS pain = 50, satisfaction VAS = 50, EQ mobility = 2, EQ pain/discomfort = 2)

Strengths and limitations

We obtained all our data from SHAR, a well-established and validated quality register. The completeness of the register and the large number of responders to PROMs both pre- and post-operative reduces confounding bias. The large number of included patients following the selection process also contributes to reducing confounding bias and heterogeneity. This gives robustness and power to our calculation and model for predicting the risk of a late re-operation using post-operative PROMs.

There are limitations to this study. Factors other than PROMs have been previously reported [10–13] as contributing to the risk of re-operation, such as high BMI and long surgery time. We have not taken all of these confounders into consideration in our model, which is a limitation that needs to be given attention when interpreting results. Therefore, before we could develop an algorithm for the detection of patients with a risk of re-operation to be used on a regular basis, such factors and others such as implant type need to be included into the model. This could be possible, but further investigations are needed. The models presented in this study only included the one-year PROMs. One could argue that the outcomes are dependent of pre-operative levels of PROM.

However, when we included pre-operative PROMs into the model the predictive power did not change (data not shown).

Conclusion

This study demonstrates that PROMs collected at one year after THR can predict future re-operations. PROMs one year following THR are stronger predictors for re-operation than age and sex. The model can estimate individualized re-operation probabilities at different time points that could be used to develop a warning system to indicate patients at high risk. Including other baseline information in the model may increase the predictive power. However, there is a delicate trade-off between model complexity and predictive power in the continued work of developing a clinical application.

Author's contributions Drafting of the manuscript was done by TE and OR. Statistical analysis and interpretation of data were done by TE, EB, SN and OR. Substantial contributions to research design, drafting, and critical revision of the manuscript were performed by TE, EB, SN, MM and OR. All authors have read and approved the final submitted manuscript.

Funding This study was supported by research grants from Dr. Felix Neuberg's foundation and Gothenburg Medical Association.

Compliance with ethical standards

Conflict of interest None of the authors have any conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors. Ethical approval was obtained from the Regional Ethical Review Board in Gothenburg, Sweden, registration number 268-17.

Informed consent Informed consent was not applicable.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Garellick G, Karrholm J, Lindahl H, Malchau H, Rogmark C, Rolfson O (2015) Swedish hip arthroplasty register annual report 2014. Göran Garellick, Gothenburg, Sweden
- Junnila M, Laaksonen I, Eskelinen A, Pulkkinen P, Ivar Havelin L, Furnes O, Marie Fenstad A, Pedersen AB, Overgaard S, Karrholm J, Garellick G, Malchau H, Makela KT (2016) Implant survival of the most common cemented total hip devices from the Nordic Arthroplasty register association database. *Acta Orthop* 87:546–553. <https://doi.org/10.1080/17453674.2016.1222804>
- Kärrholm J, Lindahl H, Malchau H, Mohaddes M, Rogmark C, Rolfson O (2016) Swedish hip arthroplasty register. Annual Report 2015. Ola Rolfson, Gothenburg, Sweden
- Espehaug B, Havelin LI, Engesaeter LB, Langeland N, Vollset SE (1998) Patient satisfaction and function after primary and revision total hip replacement. *Clin Orthop Relat Res* 351:135–48
- Lie SA, Vollset LI, Havelin ON, Furnes LB, Engesaeter SE (2004) Failure rates for 4762 revision total hip arthroplasties in the Norwegian arthroplasty register. *J Bone Joint Surg B* 86:504–509. <https://doi.org/10.1302/0301-620X.86B4.14799>
- Lübbecke A, Katz JN, Pemeger TV, Hoffmeyer P (2007) Primary and revision hip arthroplasty: 5-year outcomes and influence of age and comorbidity. *J Rheumatol* 34:394
- Adelani MA, Crook K, Barrack RL, Maloney WJ, Clohisy JC (2014) What is the prognosis of revision total hip arthroplasty in patients 55 years and younger? *Clin Orthop Relat Res* 472:1518–1525. <https://doi.org/10.1007/s11999-013-3377-9>
- Bolz KMD, Crawford RW, Donnelly B, Whitehouse SL, Graves N (2010) The cost-effectiveness of routine follow-up after primary total hip arthroplasty. *J Arthroplast* 25:191–196. <https://doi.org/10.1016/j.arth.2008.12.009>
- Sethuraman V, McGuigan J, Hozack WJ, Sharkey PF, Rothman RH (2000) Routine follow-up office visits after total joint replacement: do asymptomatic patients wish to comply? *J Arthroplast* 15:183–186. [https://doi.org/10.1016/S0883-5403\(00\)90176-1](https://doi.org/10.1016/S0883-5403(00)90176-1)
- Dy CJ, Bozic KJ, Pan TJ, Wright TM, Padgett DE, Lyman S (2014) Risk factors for early revision after total hip arthroplasty. *Arthritis Care Res* 66:907–915. <https://doi.org/10.1002/acr.22240>
- Paxton EW, Inacio MCS, Khatod M, Yue E, Funahashi T, Barber T (2015) Risk calculators predict failures of knee and hip arthroplasties: findings from a large health maintenance organization. *Clin Orthop Relat Res* 473:3965. <https://doi.org/10.1007/s11999-015-4506-4>
- Wright EA, Katz JN, Baron JA, Wright RJ, Malchau H, Mahomed N, Prokopetz JJZ, Losina E (2012) Risk factors for revision of primary total hip replacement: results from a national case-control study. *Arthritis Care Res* 64:1879–1885. <https://doi.org/10.1002/acr.21760>
- Prokopetz JJZ, Losina E, Bliss RL, Wright J, Baron JA, Katz JN (2012) Risk factors for revision of primary total hip arthroplasty: a systematic review. *BMC Musculoskelet Disord* 13:251–251. <https://doi.org/10.1186/1471-2474-13-251>
- Singh JA, Schleck C, Harmsen S, Lewallen D (2016) Clinically important improvement thresholds for Harris hip score and its ability to predict revision risk after primary total hip arthroplasty. *BMC Musculoskelet Disord* 17:256. <https://doi.org/10.1186/s12891-016-1106-8>
- Rothwell AG, Hooper GJ, Hobbs A, Frampton CM (2010) An analysis of the Oxford hip and knee scores and their relationship to early joint revision in the New Zealand joint registry. *J Bone Joint Surg Br* 92:413–418. <https://doi.org/10.1302/0301-620x.92b3.22913>
- Devane P, Horne G, Gehling D (2013) Oxford hip scores at 6 months and 5 years are associated with total hip revision within the subsequent 2 years. *Clin Orthop Relat Res* 471:3870–3874. <https://doi.org/10.1007/s11999-013-2880-3>
- Greene ME, Rolfson O, Gordon M, Garellick G, Nemes S (2015) Standard comorbidity measures do not predict patient-reported outcomes 1 year after total hip arthroplasty. *Clin Orthop Relat Res* 473:1–10
- Rolfson O, Karrholm J, Dahlberg LE, Garellick G (2011) Patient-reported outcomes in the Swedish hip arthroplasty register: results of a nationwide prospective observational study. *J Bone Joint Surg Br* 93:867–875. <https://doi.org/10.1302/0301-620x.93b7.25737>
- Greene ME, Rolfson O, Garellick G, Gordon M, Nemes S (2015) Improved statistical analysis of pre- and post-treatment patient-

- reported outcome measures (PROMs): the applicability of piecewise linear regression splines. *Qual Life Res* 24:567–573. <https://doi.org/10.1007/s11136-014-0808-3>
20. EuroQol Group (1990) EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 16(3):199–208
 21. Scott J, Huskisson E (1976) Graphic representation of pain. *Pain* 2: 175–184
 22. Devlin NJ, Krabbe PFM (2013) The development of new research methods for the valuation of EQ-5D-5L. *Eur J Health Econ* 14(Suppl 1):S1–3
 23. Heagerty PJ, Zheng Y (2005) Survival model predictive accuracy and ROC curves. *Biometrics* 61:92–105. <https://doi.org/10.1111/j.0006-341X.2005.030814.x>
 24. Shmueli G (2010) To explain or to predict? *Stat Sci* 25:289–310. <https://doi.org/10.1214/10-STS330>
 25. Pepe MS, Longton G, Leisenring W, Newcomb P, Janes H (2004) Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol* 159:882–890. <https://doi.org/10.1093/aje/kwh101>