

ARTICLE

Improving patient reported outcomes in total hip replacement and total knee replacement: What can we learn from the “best” and the “poorest” performing sites?

Thomas W. Wainwright BSc (Hons) MCSP^a, Kieran Gallagher BSc (Hons) MBBS FRCS (Orth)^b, Athanasios Pollalis MBBS^c, Tikki Immins BSc (Hons) MSc^d and Robert G. Middleton MA MBBChir FRCS (Orth)^e

a Associate Professor & Deputy Head, Orthopaedic Research Institute, Bournemouth University, Bournemouth & Orthopaedic Department, The Royal Berkshire Hospital, Reading, UK

b Visiting Fellow, Orthopaedic Research Institute, Bournemouth University, Bournemouth, UK

c Fellow, Orthopaedic Department, The Royal Berkshire Hospital, Reading, UK

d Research Development Manager, Orthopaedic Research Institute, Bournemouth University, Bournemouth, UK

e Professor; Consultant Surgeon & Head, Orthopaedic Research Institute, Bournemouth University, Bournemouth & Orthopaedic Department, The Royal Berkshire Hospital, Reading, UK

Abstract

Introduction: Oxford Hip Scores (OHS) and Oxford Knee Scores (OKS) are being collected as part of the Patient Reported Outcome Measures (PROMs) evaluation programme on total hip and knee replacement. This study compares the PROMs outcomes from best and poorest performing English hospitals, as defined by NHS England’s data of elective hip and knee surgery. This was to quantify the difference and identify the scope for improvement.

Materials and Methods: OHS and OKS were obtained from the Health and Social Care Information Centre for April 2013 to March 2014.

Results: Seven sites for OHS and 10 sites for OKS were above the upper 95% control limit. Fourteen sites for OHS and 10 sites for OKS were below the lower 95% control limit. Median pre-operative scores were similar between best and poorest performers. Median post-operative scores were 4 points higher in best performers. Top OHS-performers scored better in limping, stairs, work, transport, dressing and shopping. Top OKS-performers scored better in walking, shopping and kneeling.

Discussion: Pre-operative scores were similar for the best and poorest performers. The differences between best and poorest performing hospitals for OHS and OKS were below the minimum important difference. There was only moderate consistency for outliers. Results for any single year should be treated with caution.

Conclusions: PROMs, one of many key tools in measuring and increasing person-centered healthcare, can be useful as part of an evaluation of practice but do not always reveal the full picture. It is important that balanced measures of quality should be used when benchmarking hospitals.

Keywords

Activities of daily living (ADLs), arthroplasty, case-mix, co-morbidity, efficiency targets, Oxford Hip Score, Oxford Knee Score, patient experience, patient reported outcome measures (PROMs), performance management, person-centered healthcare, person-centered rehabilitation, socioeconomic data

Correspondance address

Ms. Tikki Immins, Orthopaedic Research Institute, Bournemouth University, 6th Floor, Executive Business Centre, 89 Holdenhurst Road, Bournemouth, BH8 8EB, UK. Email: timmins@bournemouth.ac.uk

Accepted for publication: 20 June 2017

Introduction

Oxford Hip Scores (OHS) and Oxford Knee Scores (OKS) have been collected in England as part of the Patient Reported Outcome Measures (PROMs) evaluation programme on total hip and knee replacement surgery since 2009 [1,2]. PROMs were introduced in an effort to

improve care through improving patient choice and enabling clinicians and managers to review their performance [3]. PROMs are also used by the CQC for Hospital Intelligent Monitoring [4,5] and furthermore they may be potentially linked to payment [6]. Commonly, providers deemed to be outliers are contacted by commissioners with enquiries as to what actions are to be

taken in order to improve their performance - indeed, it is proposed that in 2014/2015, providers below the lower 99.8% control limit do not receive the Best Practice Tariff (BPT) [6]. In the 2016/17 BPT proposals it is proposed that those providers below the 95% lower control limit may not receive the payments. This is a national tariff structured as part of the Department of Health's Best Practice Tariff initiative which rewards the achievement of specified standards. However, recent research suggests that the routine use and feedback of PROMs scores is yet to have much influence on provider's behaviour [7]. Hospital monitoring by the CQC uses PROMs data for elective hip and knee procedures as 2 of 57 measures of performance of a hospital trust as a whole. It is therefore vital that an understanding of the meaning of the differences is attained, to investigate the associated appropriateness.

Although the effect of patient characteristics on PROMs has been explored [8], little work has been carried out to compare providers. The aim of this study is to assess and compare outcomes of best and poorest performing hospitals for OHS and OKS and for each score's individual questions and thus to make recommendations that can stimulate improvement. We wanted to analyze if there was a real difference between the groups of hospitals and if so what can the poorest performing hospitals learn from the best performing hospitals?

Materials and Methods

Secondary analysis was carried out on provisional PROM data released on the HSCIC website (<http://www.hscic.gov.uk/proms>) for April 2013 to March 2014 for primary hip and knee replacements. These data were released in August 2014.

Patients completed the validated OHS or OKS questionnaires [9,10] before and 6 months following surgery. Each questionnaire comprises 12 multiple-choice questions which ask about the patient's experience of pain and ability to carry out normal everyday tasks such as walking, dressing and climbing stairs. Each question is scored 0 - 4 with 4 indicating little or no symptoms and 0 indicating severe symptoms. The individual question scores are then added to provide a composite score with a score of 48 being the best outcome.

We identified all the hospitals that were either above or below the 95% control limits for the case-mix adjusted health gain measurement for both OHS and OKS from April 2013 to March 2014. The case-mix adjustment is applied to the PROM scores to reflect the outcomes that would have been achieved if the hospital had been treating a national average case mix of patients. This adjusts for factors such as age, gender, co-morbidities and socioeconomic data. Sites with less than 100 procedures were first removed from the data. This resulted in 4 groups: Best performing OHS hospitals, best performing OKS hospitals, poorest performing OHS hospitals and poorest performing OKS hospitals. Patient level data were then amalgamated so that each group contained all the patient data for all the included hospitals. The median was calculated by group for the total score and also for each

question in the OHS or OKS. Since scales are ordinal, median rather than mean was used, as it is relatively unaffected by skewed distributions.

Results

Seven sites were above the upper 95% control limit for OHS and 14 sites below the lower 95% control limit. For OKS, 10 sites were above and 10 sites were below the 95% control limits. Three sites were in the poorest-performing and 2 sites were in the best performing groups for both OHS and OKS.

Table 1 Comparison of Median Oxford Hip Score and Median Oxford Knee Score for best and poorest performers

	Best performers	Poorest performers
Median Oxford Hip Score		
Pre-operative	18	17
Post-operative	44	40
Median Oxford Knee Score		
Pre-operative	19	18
Post-operative	39	35

Pre-operative OHS median scores were 18 for the best and 17 for the poorest performing groups (See Table 1). These scores improved post-operatively to 44 and 40 respectively. When analysing median scores of individual questions, the best and poorest performing groups had the same score for each question, except for pain (1 vs 0) (See Figure 1). Post-operatively the best performing group had the maximum score for all questions except dressing (putting on socks/shoes/tights), where median score was 3 (see Figure 2). The poorest performing group had the maximum score for experiencing sudden pain, pain at night, standing (up from a chair) and washing (and drying) and walking (without severe pain for 30 minutes or more) but scored a median of 3 for pain, functional abilities limping (when walking), stairs (climbing a flight of stairs) and work (does pain interfere with work?) and Activities of Daily Living (ADLs) transport (getting in and out of a car/using public transport), dressing and shopping.

For OKS, pre-operative median scores were 19 for the best and 18 for the poorest performing groups (See Table 1). Post-operatively, these scores improved to 39 and 35 respectively. Individual questions analysis showed same median scores between best and poorest performing groups for all questions, except for standing (2 vs 1) (see Figure 3). Post-operatively both groups scored a median of 4 for confidence (in knee giving way) and for washing and a median of 3 for pain, pain at night, transport, standing, limping, work and stairs (walking down a flight of stairs) (see Figure 4). The best performing sites had better median scores for walking (4 vs 3), shopping (4 vs 3) and kneeling (2 vs 1). Although the best performing sites had little improvement in kneeling (from median score of 1 to 2), the poorest performing sites had no improvement.

Figure 1 Comparison of Oxford Hip Score pre-surgery best and poorest performers

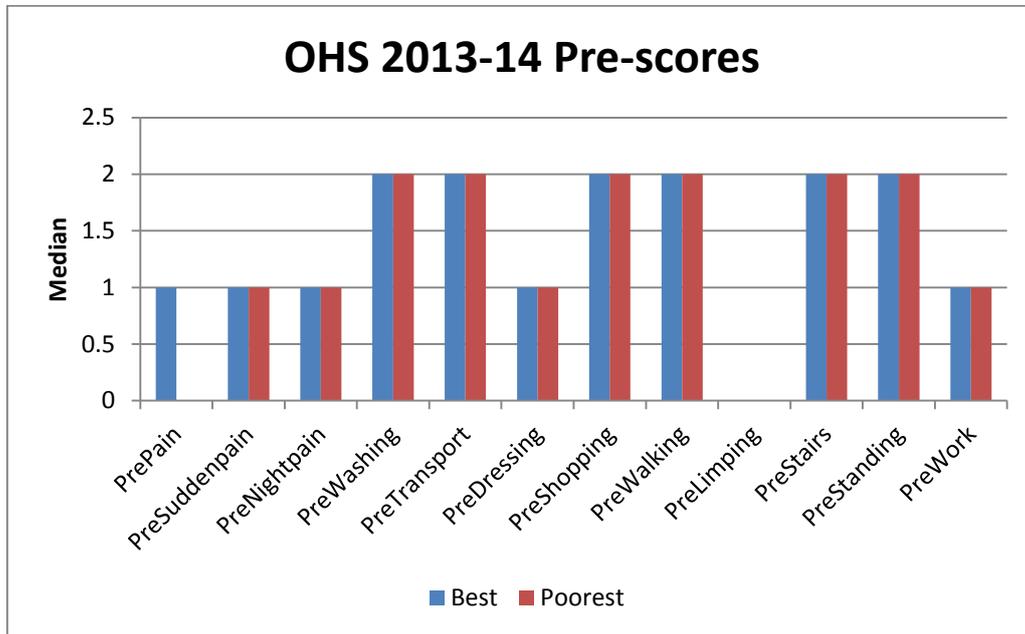


Figure 2 Comparison of Oxford Hip Score post-surgery best and poorest performers

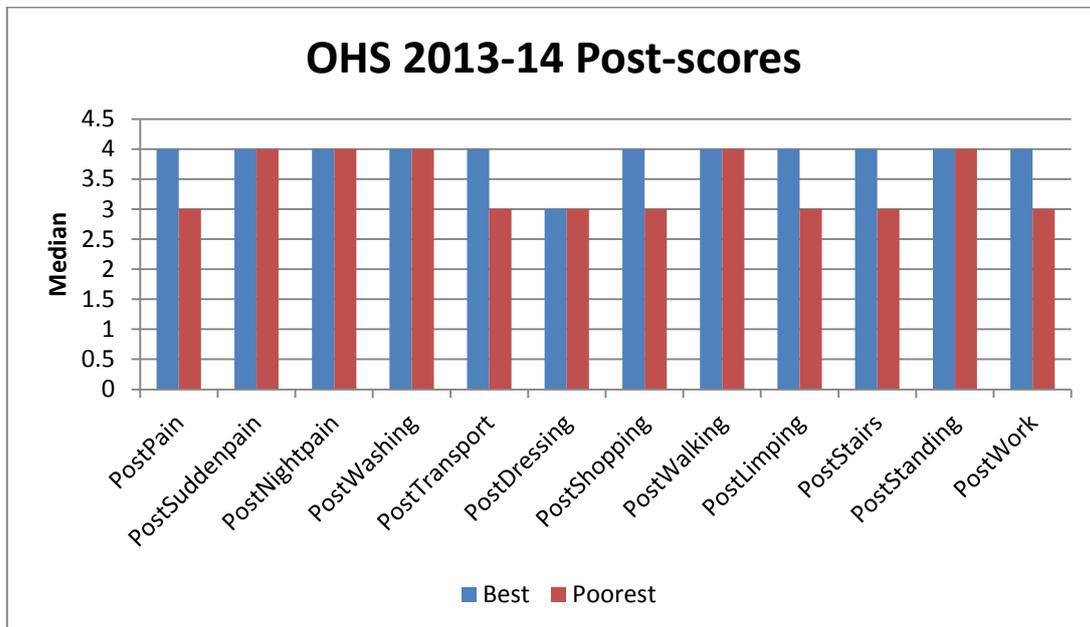


Figure 3 Comparison of Oxford Knee Score pre-surgery best and poorest performers

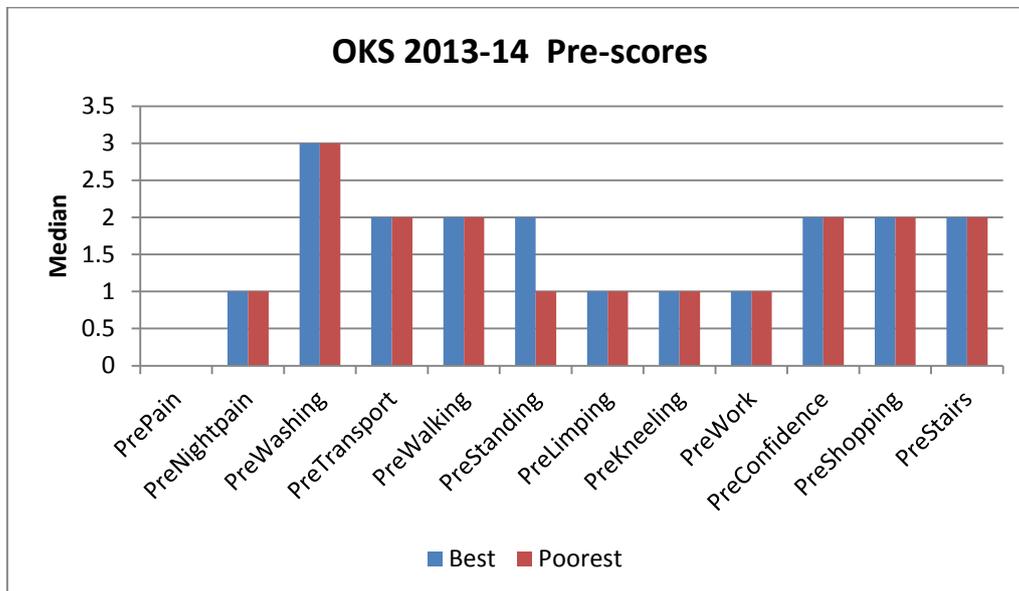
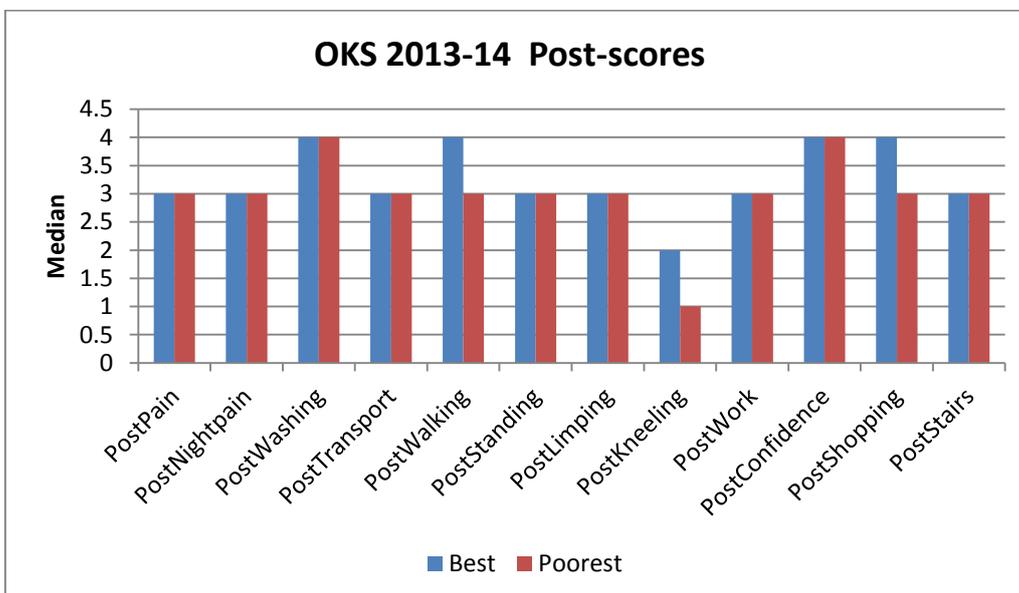


Figure 4 Comparison of Oxford Knee Score post-surgery best and poorest performers



Discussion

Pre-operative scores are considered to be the biggest determinants of outcome after a joint replacement [11,12]. Our study has shown similar pre-operative OHS and OKS in the best- and poorest-performing groups (18 vs 17 and

19 vs 18 respectively), which suggests that both groups have a similar starting point.

Research into meaningful changes for the Oxford Hip and Knee Scores have been published recently [13]. It has been suggested that the Minimally Important Difference should be used to assess the difference between groups and this is 5 points for both OHS and OKS. In our study, the

best and poorest performers displayed a difference of 4 points in their post-operative OHS (44 vs 40) and OKS (39 vs 35). Therefore, conclusions should be drawn carefully, as these changes may not demonstrate a true clinical difference. It should however be noted that both the best and poorest performers have room for improvement in the post-operative OKS.

To improve composite OHS and OKS scores it may be proposed that responses to individual questions need to improve. The best performing sites for OHS can only improve on the question of dressing. However, the poorest performing sites can introduce strategies to improve responses to pain, transport, dressing, shopping, limping, stairs and work. For OKS, both best and poorest performing sites can improve post-operative scores for pain, night pain, transport, standing, limping, work and stairs. Poorest performing sites can also improve walking and shopping. There are still considerable improvements that both groups can make to kneeling. It should be noted, however, that difficulty with answering this question has previously been reported, when patients have been told not to kneel. If this is the case, it has been suggested that patients should answer this item hypothetically [11].

When trying to reflect these results into clinical practice, factors that can influence post-operative OHS and OKS have been reported. Pre-operative OHS and OKS is consistently the most important factor amongst all available studies. Femoral component offset, age, increasing BMI and coexisting diseases can influence outcome following THR [8]. Implant brand and hospital type have been reported as determinants of outcome following TKR [14]. Jameson *et al.* [15] reported significantly smaller improvements in OHS and EQ-5D index ($P < 0.001$), and greater risk of wound complications ($P = 0.006$), re-admission ($P = 0.001$) and re-operation ($P = 0.003$) for patients with $BMI > 35.0 \text{ kg/m}^2$.

There is high-quality evidence from one large RCT to support the use of home-based physiotherapy instead of inpatient physiotherapy after primary THR or TKR surgery [16,17]. Task-oriented exercises associated with early full weight bearing improve disability (limping), pain, activities of daily living (such as walking, climbing stairs, dressing, getting in and out of car, household shopping and work-related activities) and quality of life after total hip replacement [18].

Rehabilitation programs should include exercises aimed at improving hip extensor and flexor strength and endurance to improve gait function [19]. Mizner *et al.* suggest that there is a close relationship between *quadriceps femoris* muscle group strength and the patients' ability to safely descend stairs [20]. In addition to the *quadriceps femoris* muscle group, therapeutic exercises should also address hip abductor/adductor muscle strength and incorporate incrementally greater step heights to improve dynamic lower-limb joint stability [21]. This addition may be especially important for shorter subjects and those with greater BMI. However, concerns over patient pain levels and safety, surgical incision status and prosthesis integrity usually limit early eccentric strength-training intervention [22].

Difficulties with post-operative kneeling are of major significance amongst patients undergoing total knee replacement. According to Hassaballa *et al.* [23] and Schai *et al.* [24] patients are generally uncertain about the recommendations given to them concerning kneeling and perhaps, as a result, are afraid of damaging their prosthesis. There is also a difference between the perceived and the actual kneeling ability of these patients. Encouraging patients to kneel as part of their rehabilitation programme, as well as providing clear advice and instruction on kneeling by an experienced therapist in clinic, could help in this context. Increased knee flexion is reported to improve kneeling ability, whereas a limitation in knee flexion post-operatively may be one of the factors related to greater pain or increased difficulty when kneeling [25]. Hip and ankle range of motion (ROM) can also affect kneeling and therefore should be assessed during rehabilitation.

It is, however, unrealistic to think that every patient will be able to perform all functional activities following joint replacement. Joint replacements have been designed to achieve better functional outcome during recent years [26], but it may not be possible for patients to perform demanding activities, such as kneeling, due to other joint problems or advancing years and therefore expectations of both patients and healthcare professionals should probably not be too great.

The data collection methods for the PROMs programme were derived from a multi-centre pilot study performed in 2007 [27] which recommended the use of outcome assessment at 6 months after joint replacement. The choice of a 6-month interval represented a judgment about the earliest time point in the post-operative recovery process at which the average patient has achieved all of the clinically important benefits of surgery. Although recent data support the choice of a 6-month follow-up [28], clinically important improvements in the OHS and OKS in the 6- to 12-month recovery period have been reported and therefore using both 6- and 12-month outcome data has also been recommended [29].

When analyzing the impact of disease-specific PROMs on hospital performance Varaganam *et al.* [7] reported 60% of providers performing consistently "as expected". The proportion of providers deemed to be outliers did not change over time. There was only moderate consistency in those providers deemed to be outliers for hip and knee replacement (Kappa 0.31-0.47) and although 35% of providers of hip replacement were outliers in at least one year, only 6% were consistently outliers. This lack of consistency means that results for any given provider for any given year should be treated with caution and long-term data may be needed before any action is taken. This observation, in combination with the fact of the advent of the use of PROMs for hospital monitoring, means that it is important to note that these only reflect one element of the quality of care. These should be qualified with balancing measures such as revision and readmission data. The positioning of a Hospital Trust as a poor performer has financial, operational and reputational implications. There are financial penalties for not reaching the targets for Best Practice Tariffs. Operational disruptions can arise from

inquests regarding poor performance which may not uncover valid reasons for the scores. Reputational problems can arise in the eyes of patients and staff due to national coverage of 'league tables'.

Recent research urges caution in comparison of providers' PROMs performance owing to methodological issues. Estimates of recruitment rates were found to be inaccurate as they were based on pre-operative questionnaires and hospital episodes aggregated counts, potentially biasing comparisons of providers' outcome measures [30]. Also, PROMs data are likely to have a skewed distribution and so the assumption of symmetric upper and lower control limits may be incorrect [31]. It was found that out of 237 providers of hip replacement, an extra 1.3% were classified as 'poorest performers' if symmetric control limits were used rather than simulated asymmetric control limits. The limitations of this study include the absence of age, gender and co-morbidity data of the patients. We were also unable to attain the re-admission and revision rates of the individual Hospital sites under analysis. These would allow other measures of outcome to be compared.

Conclusion

The routine use and feedback from PROMs should enable providers to review their services and instigate interventions to improve outcomes for patients as required. However, methodological challenges and achieving credible case-mix adjusted data need to be considered carefully, as current results may not always represent true clinical differences. More attention needs to be paid to how results are communicated and to the provision of advice as to what strategies may be implemented. In terms of future research, it may be worthwhile to look for strategies for improving performance on individual questions of OHS or OKS, as both providers and commissioners need more help and advice with regard to actions that need to be taken in this area.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] NHS Institute for Innovation and Improvement. Delivering Quality and Value. Focus on: Primary Hip and Knee Replacement. <http://www.qualitasconsortium.com/index.cfm/reference-material/delivering-value-quality/focus-on-hip-and-knee-replacement/>. Accessed 23 September 2016.

[2] Hildon, Z., Neuburger, J., Allwood, D., van der Meulen, J. & Black, N. (2012). Clinicians' and patients' view of metrics of change derived from patient reported outcome measures (PROMs) for comparing providers'

performance of surgery. *BMC Health Services Research* 12, 171-184.

[3] Health and Social Care Information Centre. (2016). Patient Reported Outcome Measures (PROMs) in England - A guide to PROMs methodology 2016. http://digital.nhs.uk/media/1537/A-Guide-to-PROMs-Methodology/pdf/PROMs_Guide_V8.pdf. Accessed 23 September 2016.

[4] Care Quality Commission. (2015). Monitoring NHS acute hospitals. Available at: <http://www.cqc.org.uk/public/hospital-intelligent-monitoring>. Accessed 23 September 2016.

[5] Institute for Innovation and Improvement. Commissioning for Quality and Innovation. Guidance for 2015/16. Available at: <https://www.england.nhs.uk/wp-content/uploads/2015/03/9-cquin-guid-2015-16.pdf>. Accessed 23 September 2016.

[6] Monitor and NHS England. NHS National Tariff Payment System 2016/17. Available at: <https://www.gov.uk/government/publications/nhs-national-tariff-payment-system-2016/17>. Accessed 23 September 2016.

[7] Varaganam, M., Hutchings, A., Neuburger, J. & Black, N. (2014). Impact on hospital performance of introducing routine patient reported outcome measures in surgery. *Journal of Health Services Research & Policy* 19 (2) 77-84.

[8] Judge, A., Arken, N.K., Batra, R.N., Thomas, G., Beard, D., Javaid, M.K., Cooper, C., Murray, D. & Exeter Primary Outcomes Study (EPOS) Group. (2013). The association of patient characteristics and surgical variables on symptoms of pain and function over 5 years following primary hip-replacement surgery: a prospective cohort study. *BMJ Open* 3 (3) pii: e002453.

[9] Dawson, J., Fitzpatrick, R., Carr, A. & Murray, D. (1996). Questionnaire on the perceptions of patients about total hip replacement. *The Journal of Bone and Joint Surgery. British Volume* 78 (2) 185-190.

[10] Dawson, J., Fitzpatrick, R., Murray, D. & Carr, A. (1998). Questionnaire on the perceptions of patients about total knee replacement. *The Journal of Bone and Joint Surgery. British Volume* 80 (1) 63-69.

[11] Murray, D.W., Fitzpatrick, R., Rogers, K., Pandit, H., Beard, D.J., Carr, A.J. & Dawson, J. (2007). The use of the Oxford hip and knee scores *The Journal of Bone and Joint Surgery. British Volume* 89 (8) 1010-1014.

[12] Lim, J.T., Luscombe, K.L., Jones, P.W. & White, S.H. (2006). The effect of preoperative symptom severity on functional outcome of total knee replacement--patients with the lowest preoperative scores achieve the lowest marks. *Knee* 13 (3) 216-219.

[13] Beard, D.J., Harris, K., Dawson, J., Doll, H., Murray, D.W., Carr, A.J. & Price, A.J. (2015). Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. *Journal of Clinical Epidemiology* 68 (1) 73-79.

[14] Baker, P.N., Deehan, D.J., Lees, D., Jameson, S., Avery, P.J., Gregg, P.J. & Reed, M.R. (2012). The effect of surgical factors on early patient-reported outcome measures (PROMs) following total knee replacement. *The Journal of Bone and Joint Surgery. British Volume* 94 (8) 1058-1066.

- [15] Jameson, S.S., Mason, J.M., Baker, P.N., Elson, D.W., Deehan, D.J. & Reed, M.R. (2014). The impact of body mass index on patient reported outcome measures (PROMs) and complications following primary hip arthroplasty. *Journal of Arthroplasty* 29 (10) 1889-1898.
- [16] Health Quality Ontario. (2005). Physiotherapy rehabilitation after total knee or hip replacement: an evidence-based analysis. *Ontario Health Technology Assessment Series* 5 (8) 1-91.
- [17] Mahomed, N., Wright, J.G., Hawker, G., Davis, A.M. & Badley, E.M. (2004). Randomized controlled trial of home based vs. inpatient rehab following total joint arthroplasty: functional outcomes and patient satisfaction [abstract] The Canadian Orthopaedic Association 59th Annual Meeting. 2004 Jun 17-20; Calgary, AB.
- [18] Monticone, M., Ambrosini, E., Rocca, B., Lorenzon, C., Ferrante, S. & Zatti, G. (2014). Task-oriented exercises and early full weight-bearing contribute to improving disability after total hip replacement: a randomized controlled trial. *Clinical Rehabilitation* 28, 658-668.
- [19] Horstmann, T., Listringhaus, R., Brauner, T., Grau, S. & Mündermann, A. (2013). Minimizing preoperative and postoperative limping in patients after total hip arthroplasty: relevance of hip muscle strength and endurance. *American Journal of Physical Medicine & Rehabilitation* 92 (12) 1060-1069.
- [20] Mizner, R.L., Petterson, S.C. & Snyder-Mackler, L. (2005). Quadriceps strength and the time course of functional recovery after total knee arthroplasty. *Journal of Orthopaedic & Sports Physical Therapy* 35 (7) 424-436.
- [21] Mian, O.S., Thom, J.M., Narici, M.V. & Baltzopoulos, V. (2007). Kinematics of stair descent in young and older adults and the impact of exercise training. *Gait Posture* 25 (1) 9-17.
- [22] Nyland, J., Frost, K., Quesada, P., Angeli, C., Swank, A., Topp, R. & Malkani, A.L. (2007). Self-reported chair-rise ability relates to stair-climbing readiness of total knee arthroplasty patients: A pilot study. *Journal of Rehabilitation Research & Development* 44 (5) 751-759.
- [23] Hassaballa, M.A., Porteous, A.J. & Newman, J.H. (2004). Observed kneeling ability after total, unicompartmental and patellofemoral knee arthroplasty: perception versus reality. *Knee Surgery, Sports Traumatology, Arthroscopy* 12, 136-139.
- [24] Schai, P.A., Gibson, A.J. & Scott, R.D. (2001). Kneeling ability after total knee arthroplasty. *Clinical Orthopaedics and Related Research* 367, 195-200.
- [25] Mulholland, S.J. & Wyss, U.P. (2001). Activities of daily living in non-Western cultures: range of motion requirements for hip and knee joint implants. *International Journal of Rehabilitation Research* 24, 191-198.
- [26] Nicholls, M.A., Selby, J.B. & Hartford, J.M. (2002). Athletic activity after total joint replacement. *Orthopaedics* 25, 1283-1287.
- [27] Browne, J., Jamieson, L., Lewsey, J., van der Meulen, J., Black, N. & Cairns, J. (2007). Patient Reported Outcome Measures (PROMs) in Elective Surgery, Report to the Department of Health, UK. 2007. Available at: https://www.lshtm.ac.uk/php/departmentofhealthservicesresearchandpolicy/assets/proms_report_12_dec_07.pdf. Accessed 23 September 2016.
- [28] Judge, A., Arden, N.K., Kiran, A., Price, A., Javaid, M.K., Beard, D., Murray, D. & Field, R.E. (2012). Interpretation of patient-reported outcomes for hip and knee replacement surgery: identification of thresholds associated with satisfaction with surgery. *The Journal of Bone and Joint Surgery. British Volume* 94 (3) 412-418.
- [29] Browne, J.P., Bastaki, H. & Dawson, J. (2013). What is the optimal time point to assess patient-reported recovery after hip and knee replacement? a systematic review and analysis of routinely reported outcome data from the English patient-reported outcome measures programme. *Health and Quality of Life Outcomes* 11, 128.
- [30] Hutchings, A., Neuburger, J., van der Meulen, J. & Black, N. (2014). Estimating recruitment rates for routine use of patient reported outcome measures and the impact on provider comparisons. *BMC Health Services Research* 14 (1) 1-17.
- [31] Neuburger, J., Cromwell, D.A., Hutchings, A., Black, N. & van der Meulen, J.H. (2011). Funnel plots for comparing provider performance based on patient-reported outcome measures. *BMJ Quality & Safety* 20, 1020-1026.