Predictors of clinical outcome in total hip and knee replacement : a methodological appraisal of implants and patient factors

Keurentjes, J.C.

10

Discussion

Thesis Summary

Implants In chapter 2 (p. 9), we systematically searched and appraised the current literature, regarding the probability of revision surgery at ten years for each individual Total Hip Implant. We compared the study results to the NICE benchmark,[35] and found that 8 out of 34 acetabular cups and 15 out of 32 femoral stems outperform the benchmarks. 16 out of 34 acetabular cups and 6 out of 32 femoral stems performed significantly worse than the NICE benchmark. Most studies were of low methodological quality, the risk of bias is therefore high.

Patient and Surgeon Factors In this thesis, we investigated whether two patient characteristics, namely the patients Socio-Economic Position and the preoperative radiographic severity were predictors of improvement in HRQoL and patient satisfaction after THR/TKR.

In chapter 3 (p. 37), we questioned whether more disadvantaged Socio-Economic Position is associated with an lower improvement in Health-Related Quality of Life (HRQoL) and a lower patient satisfaction after THR/TKR in a multi-center cohort study. We found no differences in HRQoL improvement in THR patients and small, clinically irrelevant differences in HRQoL improvement in some subscales for TKR patients. Additionally, we found no differences in patient satisfaction, both for THR patients and TKR patients.

In conclusion, Socio-Economic Position is no useful patient characteristic to predict HRQoL improvement and patient satisfaction in the Netherlands.

In chapter 4 (p. 55), we assessed whether the pre-operative radiographic OA severity is related to the improvement in HRQoL after THR or TKR, both at the population and individual level. Severe OA patients improved more and had a higher probability of a relevant improvement in physical functioning after both THR and TKR. Patient satisfaction was also higher in severe OA TKR patients. In conclusion, the radiographic OA severity could be a useful patient characteristic to predict HRQoL improvement and patient satisfaction.

Research Methodology

Competing Risks: In chapter 5 (p. 71), we assessed how much bias is introduced in the estimation of the probability of revision surgery, when a crucial assumption of the Kaplan-Meier estimator is violated. Independence of the time to event and the censoring distribution is assumed in the Kaplan-Meier estimator. In the presence of competing events, this assumption does not hold. Using the Kaplan-Meier estimator when competing risks are present, will *always* lead to an overestimation of the cumulative probability in question.

Clinimetrics: In chapter 6 (p. 81), we aimed to summarise minimal clinically important differences (MCIDs) after total hip (THR) or knee replacement (TKR) in health-related quality of life (HRQoL), measured using the Short-Form 36 (SF-36). We also aimed to improve the precision of MCID estimates by means of meta-analysis. Our systematic review of the literature yielded three studies, each describing a distinct study population: primary THR, primary TKR and revision THR. No synthesis of study results can be given. The MCID estimates which we have found were not validated using external criteria and had limited precision. Nonetheless, these are the best known estimates of MCIDs in HRQoL after THR and TKR to date. We therefore advise cautious use of these MCIDs as absolute thresholds.

In chapter 7 (p. 93), we aimed to determine Clinically Important Differences (CIDs) in Health-Related Quality of Life (HRQoL) after Total Hip (THR) or Total Knee Replacement

(TKR), using the Short-Form 36 (SF36). CIDs are more relevant than MCIDs in THR and TKR, as one would expect a substantial improvement in HRQoL after joint replacement, instead of just a minimal improvement. CIDs of Physical Functioning, Role Physical, Bodily Pain and Social Functioning were validated by the validation question "knowing what your hip or knee replacement surgery did for you, would you still have undergone this surgery (yes / no)?". CIDs of all other subscales should be used cautiously, as these were not validated using external criteria.

In chapter 8 (p. 105), we aimed to define Patient Acceptable Symptom State (PASS) thresholds for the Oxford Hip Score (OHS) and Oxford Knee Score (OKS) at mid-term follow-up. Receiver operating characteristic (ROC) curves identified a PASS threshold of 42 points for the OHS after THR and 37 points for the OKS after TKR. THR patients with an OHS $\geq$ 42 points and TKR patients with an OKS $\geq$ 37 points had a higher Numeric Rating Scale of Satisfaction and a larger odds of being willing to undergo surgery again. However, PASS thresholds differed considerably between relevant subgroups. PASS thresholds appear larger at mid-term follow-up than at 6 months after surgery. Without external validation, we would advise against using these PASS thresholds as absolute thresholds in defining whether or not a patients has attained an acceptable symptom state after THR/TKR.

Questionnaire Mode Preference: In chapter 9 (p. 119), we assessed patient preference for the questionnaire mode in a multi-center cohort study. The majority of THR and TKR patients prefer pen-and-paper questionnaires. Patients preferring electronic questionnaires differed from patients who preferred pen-and-paper questionnaires. Restricting the mode of PROMs to electronic questionnaires might introduce selection bias.

Recommendations for Future Research

Implants The ideal THI has a low probability of revision surgery and has little systemic adverse effects. In chapter 2 (p. 9), we systematically searched and appraised the literature, regarding the probability of revision surgery of THI at ten years follow-up. A

number of THI outperform current benchmarks. However, the majority of found survival estimates were based on a single study, performed in a single center. The results of this study should not be viewed as conclusive evidence, but as the best available evidence at *this* point in time.

Post-market surveillance studies, which are summarised in our systematic review (chapter 2 (p. 9)), remain of vital importance to detect implants with a high probability of revision surgery. A major drawback of such studies is that, if published at all, their results are available many years after an implant has been introduced in clinical practice. National joint registries, given an annual update, solve the first problem, but not the second. Solely relying on post-market surveillance and national joint registry studies to detect poor implants will therefore expose many patients to unproven designs and facilitates large-scale implant recalls, as we have seen in the ASR case.[111]

Imaging techniques, such as 3D Röntgen Stereophotogrammetry Analysis (RSA), could play a crucial role in preventing future implant disasters. The probability of revision surgery can be predicted using RSA.[42, 114, 235] A recent study shows that implants, which have published RSA studies at two years follow-up, have 22–35% less revisions up to 5 years after surgery.[119] Phased introduction of new implants using RSA could therefore lead to better patient care and could substantially reduce health-care costs associated with revision surgery.

Patient and Surgeon Factors We have assessed the role of the patients Socio-Economic Position and the severity of preoperative radiographic OA in predicting the Patient-Reported Outcome after THR and TKR. The number of potential predictors of the improvement in clinical outcome is endless.

One potential predictor, which is often stressed in the literature, is the preoperative patient expectation of the outcome after joint replacement.[166] These expectations can be modified by preoperative educational classes, which opens up possibilities for preoperative optimalisation.[236] In future cohort studies, the role of pre-operative expectations on the probability of a relevant improvement in HRQoL should be studied, in order to investigate whether or not expectation management can lower the relatively

high rate of dissatisfaction after joint replacement.

Research Methodology

Competing Risks: In chapter 5 (p. 71), we have shown that the Kaplan-Meier (KM) estimator introduces bias in the presence of competing events. When estimating the cumulative probability of revision surgery, competing events are likely to occur in the case of THR or TKR. A recently developed guideline for the statistical analysis of arthroplasty data acknowledges that the KM estimator yields biased results.[237] Unfortunately, the authors of this guideline miss the point in interpreting the consequences of this bias. Two poor arguments in favour of the KM estimator are proposed. In the first place, "Is the difference (i.e. the amount of bias)... clinically important?" We have shown that the amount of bias depends on the number of competing events (i.e. the number of patients who have died), compared to the number of events of interest (i.e. the number of patients who have undergone revision surgery). However, why would one be willing to accept *any* form of bias, especially when it is possible to eliminate such bias using freely available tools? In the second place, the authors state that "The KM estimates of implant failure are more clinically meaningful and straightforward to interpret for clinicians and patients". This argumentation is flawed, since both the KM estimator and the cause-specific cumulative incidence estimator estimate the cumulative probability of a certain event as a function of time. Thus, the clinical meaning of the KM estimator and the cumulative incidence estimator is identical. The only difference is that the cumulative probability of being event-free is presented by the KM estimator, while the cumulative incidence estimator presents the cumulative probability of having the event of interest. Surely, getting accustomed to this slightly different way of presenting the probability of revision surgery is worthwhile, as it permits unbiased estimation of the outcome of interest.

Clinimetrics: In this thesis we have summarised the literature regarding MCIDs in HRQoL after THR and TKR, we have estimated CIDs in HRQoL after THR and TKR using an innovative approach and we have estimated PASS in Joint-specific Patient Reported Outcome Measures.

A number of issues remain to be addressed. To date, no CIDs have been established for neither OHS nor OKS, two often used joint specific Patient Reported Outcome Measures. In esstimating these CIDs, future studies could compare our innovative approach to the approach of Chesworth et al.[203]

Questionnaire Mode Preference: In chapter 9 (p. 119), we found that the vast majority of THR and TKR patients prefer pen-and-paper questionnaires, when participating in a cohort study on the improvement in Health-Related Quality of Life after THR or TKR.

In the past few years, tablet computers have gained in popularity. Recent evidence suggests that the acceptance and satisfaction rates of tablet computers are high amongst senior users.[238] It would be interesting to see whether or not the rising popularity of tablet computers will affect the patient preference for pen-and-paper questionnaires in the near future.