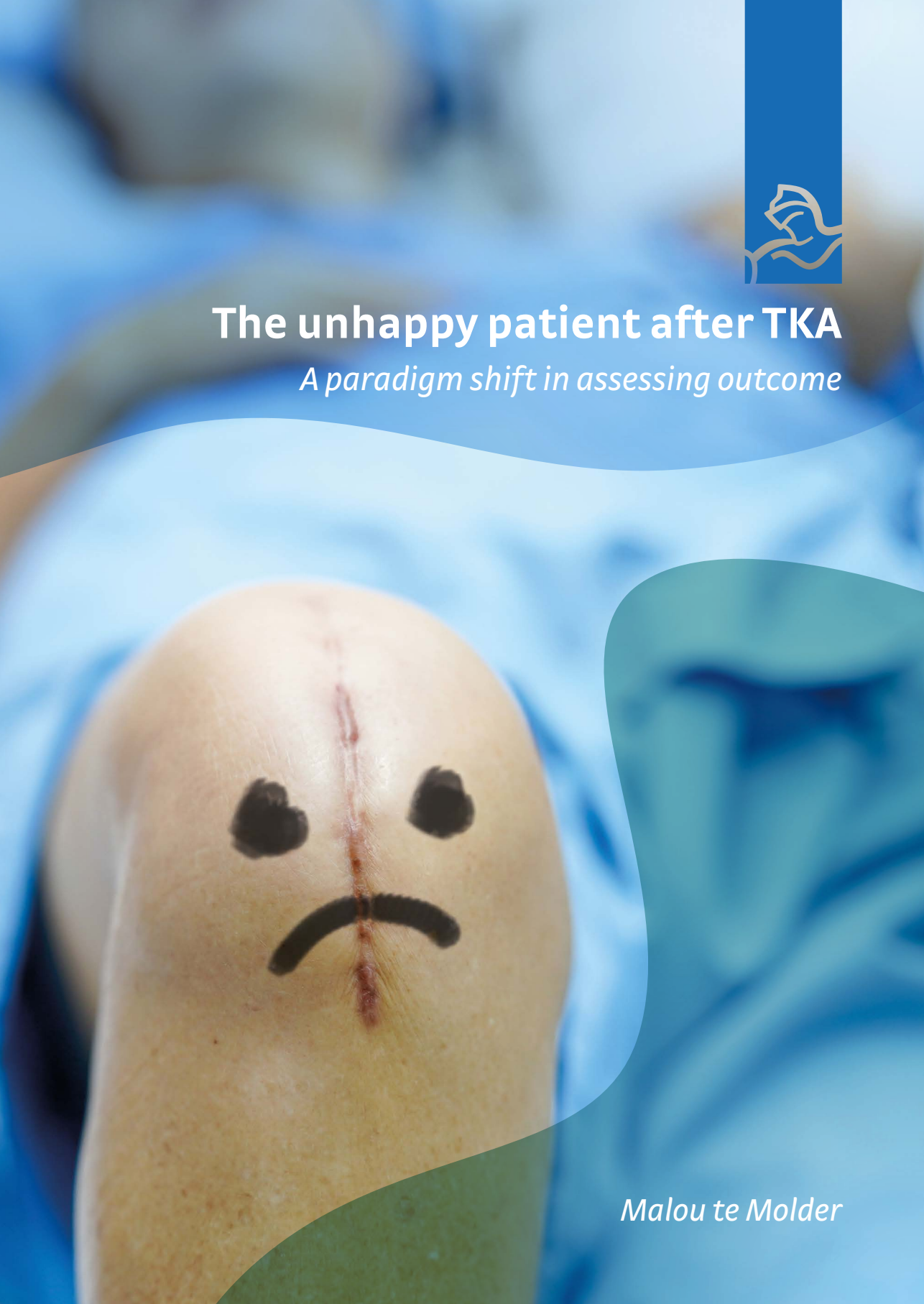
The unhappy patient after TKA

A paradigm shift in assessing outcome



Malou te Molder

The unhappy patient after TKA

A paradigm shift in assessing outcome

Malou te Molder

The unhappy patient after TKA

A paradigm shift in assessing outcome

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op
donderdag 25 april 2024
om 10.30 uur precies

door

Malou Elsa Maria te Molder
geboren op 8 mei 1991
te Winterswijk

Colofon

Cover design and layout: burorub grafisch ontwerp, Nijmegen.
Printing: Koninklijke Van der Most BV, Heerde.

Publication of this thesis was financially supported by the Sint Maartenskliniek, Nijmegen.

A number of studies described in this thesis were supported by a grant from the LROI foundation (LROI RG 2020-003).

The work presented in this thesis was carried out within the Radboud Institute for Health Sciences.

Promotoren:
Prof. dr. M. de Kleuver
Dr. C.H.M. van den Ende

Copromotoren:
Dr. P.J.C. Heesterbeek (Sint Maartenskliniek)
Dr. J.M.H. Smolders (Sint Maartenskliniek)

Manuscriptcommissie:
Prof. dr. K.C.P. Vissers
Prof. dr. H.J. Schers
Prof. dr. H. Vandenneucker (Katholieke Universiteit & Universitair Ziekenhuis Leuven, België)

Table of contents

Chapter 1

General introduction

Introduction

This thesis focuses on evaluating and defining the concept of poor response after total knee arthroplasty (TKA). The term poor response will be used in this thesis as a comprehensive concept.

Prevalence and incidence of total knee arthroplasty

Osteoarthritis (OA) of the knee is a common degenerative joint disease, causing pain, stiffness, functional impairments and a reduction of quality of life.[1] The number of people with knee OA is expected to increase in the coming decades due to ageing of the population and the obesity pandemic.[2] Since no cure is available for knee OA, core elements of treatment encompasses conservative treatment options (e.g. education, pain medication and exercise therapy). In case of failure of conservative treatment options surgical treatment with TKA is considered.[3] (Inter)national guidelines specify the following indications for a TKA procedure in patients with end-stage knee OA: inadequate pain relief from conservative treatment, function loss and radiological evidence of knee OA.[4] 96% of the registered TKAs in the Netherlands are performed due to OA.[5]

A TKA is a successful and cost-effective surgical treatment. Especially when looking at objective medical outcomes, i.e. survival and performance of prostheses, and revision rates.[6] In 2019, 25.885 primary TKA procedures were performed in the Netherlands[7], and an absolute increase to 57.893 procedures is forecasted.[8] Worldwide an average of around 135 TKAs per 100,000 persons are performed[9], and is projected to increase continuously over the next decade to 299 per 100,000 persons.[10] A considerable number of patients experience poor outcome after TKA, typically assessed through their dissatisfaction with the replaced knee.[11-13] The ratio of 1 dissatisfied patient for every 5 TKAs has persisted over the last decade, despite improvements in surgical technologies and treatments.[12,14] It is worth noting that a recent study conducted a comprehensive review of existing literature and found that the current rate of patient dissatisfaction following TKA[15] stands at 10%. This finding differs from previous studies, which might be attributed to the latter study's utilization of a more stringent cut-off point. Nevertheless, a high proportion of patients are dissatisfied with their TKA due to insufficient pain relief, limitations in physical function, complications, unmet expectations and diminished health-related quality of life.[11-13,16] Dissatisfied patients incurred greater overall societal costs than those who were satisfied during the first year after TKA.[17-19] This difference in cost was primarily attributed to higher indirect costs, particularly the duration of time patients and their healthcare providers were unable to engage in paid employment. This suggests that dissatisfied patients experience greater delays in return to work, reduced productivity, and may require a greater amount of healthcare provider support with activities of daily living.[19] Therefore, the massive projected growth in demand for TKA will inevitably place an immense burden on the future healthcare systems across the world, with increasing costs and limited resources.

Reporting of patients' dissatisfaction with TKA

For a long time, the concept of patient dissatisfaction has received little attention compared to patient satisfaction since it has been assumed to be the opposite of satisfaction and thus already defined.[20] Therefore a series of assumptions have been made about dissatisfaction, which may or may not compromise its validity or usefulness. However, over the years, much literature suggested that the satisfaction-dissatisfaction linearity is unfounded.[20]

In 1993, Verkuisen described dissatisfaction as follows:

"It is a subjective transformation, which is complicated and involves the crystallization of a strong, undifferentiated, vague, negative emotion experienced immediately (after an untoward event/experience) into a more stable negative interpretation of the experience."[20]

This statement underlines that patient dissatisfaction is not a well-defined outcome measure. Dissatisfaction with the outcome of TKA is subjective and expected to depend on a combination of several factors, weighted according to the demands and expectations of each individual patient.[13,21] Identification of the causes of patient dissatisfaction is important to be able to improve patient selection for TKA, and also to assist dissatisfied patients with their specific problems.[16] For orthopedic surgeons and other health care professionals involved in the care of TKA patients, it is an ongoing challenge to understand and therefore help dissatisfied patients.[16] The available orthopedic literature on patient dissatisfaction is somewhat limited, given the numerous ways in which dissatisfaction can be measured. The wide range of reported dissatisfaction rates can be attributed, in part, to the diversity in answering options, wording, and timeframe used across studies. It is important to acknowledge that this variability in measurement methods contributes to the challenges in accurately assessing and comparing dissatisfaction rates among different studies. Despite previous work on identifying causes of dissatisfaction, there remains a sense that there is not sufficient understanding of the mechanism of dissatisfaction after TKA to effectively address dissatisfaction, and there remains a call to improve this understanding.

Evaluation of TKA outcomes

In the field of TKA, outcome research has evolved to reflect the continued developments in surgical techniques and prosthetic designs.[22] Traditionally survival analysis was employed to ascertain the longevity of TKAs.[22] However, the occurrence of early failure in TKAs has become a rare complication, with up to 96% of TKAs in situ after 10 years.[22,23] Due to these improvements in survivorship rates, the focus of outcome research has shifted towards evaluating the impact of TKA on patients, specifically focusing on the reduction of knee pain and restoring functional abilities.[13,22,24,25] Clinician-based outcome measures, such as the Knee Society Clinical Rating System[26] represent the early shift from focusing on the survivorship of the knee prosthesis towards the patient.[22] However, previous research shows that a successful outcome according to the physician is not a guarantee for treatment success as perceived by patients.[27,28]

From focus on the "happy" to focus on the "unhappy" TKA: a paradigm shift

For years, the evolution and continued developments of surgical techniques and prosthetic designs in TKA have been focused on optimizing outcomes in patients who are already performing well on group level. This is also reflected by the ceiling effects that were shown for many patient-reported outcome measures (PROMs) used for the assessment of patient improvement following TKA.[29] However, it is difficult to further identify areas for improvement when focusing on the majority who do well. The concept of the law of diminishing returns has brought about a paradigm shift in the field, shifting the focus from patients who have successful outcomes after TKA to those who experience poor outcomes. This shift recognizes the importance of addressing the needs and improving the outcomes for patients who do not achieve optimal results from the procedure. By prioritizing patient with poorer outcomes, healthcare providers and researchers can direct their efforts towards enhancing their overall

experience and functional recovery after TKA. Age-specific PROMs reflect this new way of thinking as they allow for the identification of poor outcomes within specific age groups, for example younger patient with an active lifestyle, who have higher expectations of the outcome after the TKA procedure.[30,31]

Common language and rationale behind a uniform concept of poor response after TKA

Currently, there appears to be no consensus regarding which uniform/shared vocabulary is most appropriate for defining the phenomenon of poor outcome after TKA. The heterogeneity and lack of uniformity in outcome measures applied to TKA confirms this and thus hinder the ability to understand and improve TKA outcomes. International need for a multidimensional combination of outcome domains (e.g. pain and function) has been recognized to describe poor outcome after TKA, which is necessary to identify the patients and patient groups in need for improvement of outcome.[32] The ultimate goal of identifying poor response after TKA is to minimize the proportion of patients who experience a poor outcome. Furthermore, it provides opportunities for both clinical care and research, because it can facilitate: 1) comparisons of poor response after TKA over time and across studies, hospitals and countries for transparency and quality improvement; 2) a solid basis for continuous outcome monitoring to reduce undesirable practice variation due to surgeons and hospital related characteristics; and 3) identification of patients with poor response after TKA to examine the predictive value of preoperative factors in both research and clinical practice. During the research process of this thesis, the researchers had an ongoing discussion on the vocabulary, terminology or concepts that best reflects the phenomenon of this thesis. Terminology that come together in this thesis are: failure, poor outcome, dissatisfaction, non-responder, adverse consequences of TKA, not happy and unhappy. Eventually, the researchers consolidated various terminologies into a comprehensive concept called "poor response to TKA". Different terminology describing more and less the same phenomenon is demonstrated in Figure 1.

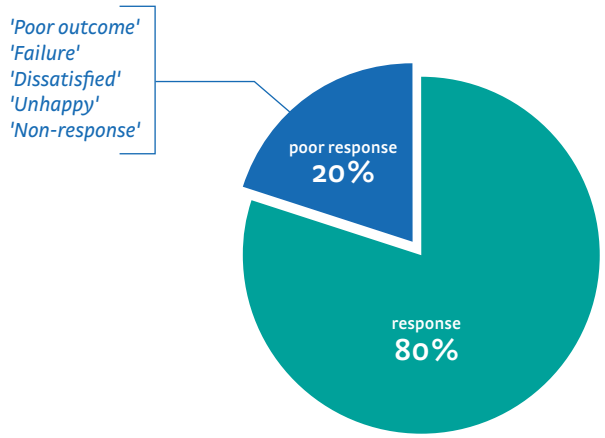


Figure 1. Different terminology describing more and less the same phenomenon

How to measure poor response after TKA

A definition of poor response after TKA should contain multidimensional criteria to identify patients with a poor outcome, and should state the outcome domain(s) and corresponding type of threshold(s) (e.g. absolute/relative, change/cut-off), value(s) and moment of follow-up. Various approaches have been proposed in the literature for a meaningful interpretation of PROMs, as outlined in Table 1. These approaches can be utilized independently or in combination when defining poor response to TKA. However, it is important to note that poor response and response to TKA cannot be simply classified as opposing states, because many experiences after TKA can lead patients to consider the outcome of their TKA as favorable or unfavorable, and the relative importance of each experience may vary with the individual patient. Moreover, outcome domains or combinations of outcome domains contributing to the perception of poor response after TKA may differ from one patient to another, and need to be explored.

Table 1. Different approaches for a meaningful interpretation of PROMs

Approach	Description
Absolute cut-off	Represents the absolute cut-off for follow-up scores.
Absolute change	Represents the absolute change from baseline to follow-up.
Relative change	Represent the relative change from baseline to follow-up.
Minimal Important Change (MIC)	Represents the smallest change in score that patients consider important.[33]
Minimal Detectable Change (MDC)	Represents the smallest change in score that can be detected by the instrument beyond measurement error.[34]
Patient Acceptable Symptom State (PASS)	Represents the lowest cut-off value for the health status that the average patient considers acceptable.[35,36]

Abbreviations: PROMs, patient-reported outcome measures

So far, it is unknown which type, amount and combination of outcome domains, type of threshold, value(s) and moment of follow-up should be used in a definition for measuring poor response after TKA. Therefore, this thesis will explore which outcome domains and interpretations should be incorporated in a definition of poor response after TKA.

Aims of the thesis

The purpose of this thesis is to gain more in depth insight into patients' and knee specialists' perspectives on poor response after TKA. The thesis is guided by six research questions:

1. How can levels of activity and participation after TKA quantitatively be measured? To answer this question, the Oxford Knee Score – Activity and Participation Questionnaire (OKS-APQ) was translated into Dutch and the measurement properties were evaluated (Chapter 2).

2. What definitions of poor response after TKA are reported in the literature? To answer this question, the literature was reviewed and definitions of poor response after TKA were summarized (Chapter 3).

3. What are adverse consequences of TKA? This question was studied in a qualitative study in which 25 patients and 15 knee specialists were interviewed (Chapter 4).

4. What is the prevalence, discriminative accuracy (sensitivity, specificity, predictive value and Youden index) and overlap of existing and newly developed definitions of poor response after TKA? To answer this research question, data of two large databases were used to compare the prevalence, discriminative accuracy and overlap of different definitions of poor response after TKA (Chapter 5).

5. How are definitions for poor response after TKA ranked by a panel of international experts? (chapter 5). Based on results of the previous studies, an international, three-round, online modified Delphi study was conducted with fifty-one panelists (Chapter 6).

6. How are adverse consequences after TKA prioritized by postoperative patients and knee specialists? Differences between what patients consider important consequences and what knee specialists think patients consider to be important are described (Chapter 7).

In Chapter 8, the main findings of this thesis are discussed. Furthermore, implications for clinical practice and recommendations for future research are given.

Workflow of this thesis

As described above, the conceptualization of poor response to TKA is complex and requires an in-depth understanding of the outcome domains and interpretations incorporated in a definition of poor response after TKA. In this thesis the following steps were taken to gain insight into patients' and knee specialists' perspectives on poor response to TKA as shown in figure 2.

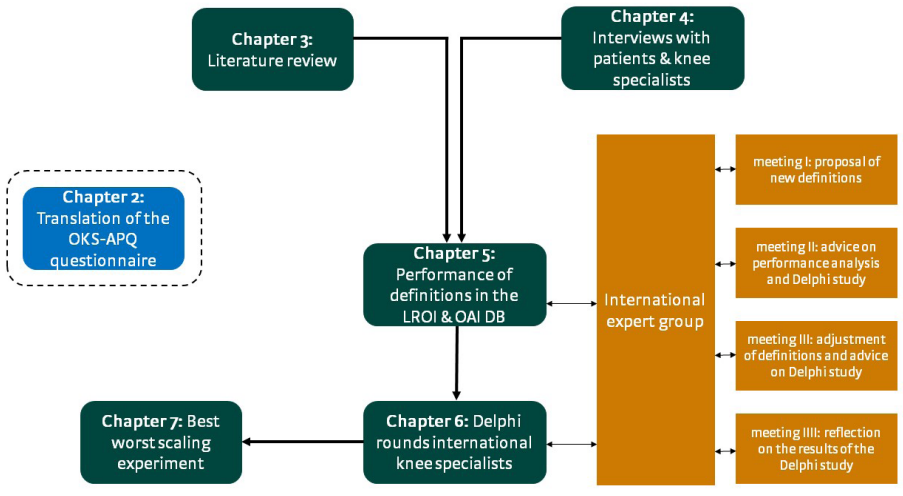


Figure 2. Schematic illustration of the outline of this thesis

References

1. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. Lancet 2011;377(9783):2115-26.
2. Neogi T, Zhang Y. Epidemiology of osteoarthritis. Rheum Dis Clin North Am 2013;39(1):1-19.
3. Zhang W, Moskowitz RW, Nuki G, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthritis Cartilage 2008;16(2):137-62.
4. Gademan MG, Hofstede SN, Vliet Vlieland TP, Nelissen RG, Marang-van de Mheen PJ. Indication criteria for total hip or knee arthroplasty in osteoarthritis: a state-of-the-science overview. BMC Musculoskelet Disord 2016;17(1):463.
5. (LROI) DAR. Patient characteristics of all patients with a registered primary knee arthroplasty by type of diagnosis in the Netherlands in 2021. 2021. Retrieved from: https://www.lroi-report.nl/knee/patient-characteristics/. Date last accessed: 19-06-2023.
6. Evans JT, Walker RW, Evans JP, Blom AW, Sayers A, Whitehouse MR. How long does a knee replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. Lancet 2019;393(10172):655-663.
7. (LROI) DAR. Number of registered knee arthroplasties per year of surgery (2007-2021) in the LROI in May 2022. 2022. Retrieved from: https://www.lroi-report.nl/knee/numbers/. Date last accessed: 05-04-2023.
8. Kuijer PPFM, Burdorf A. Prevention at work needed to curb the worldwide strong increase in knee replacement surgery for working-age osteoarthritis patients. Scand J Work Environ Health 2020;46(5):457-460.
9. OECD. Health at a Glance. 2019. OECD Indicators.
10. Klug A, Gramlich Y, Rudert M, et al. The projected volume of primary and revision total knee arthroplasty will place an immense burden on future health care systems over the next 30 years. Knee Surg Sports Traumatol Arthrosc 2021;29(10):3287-3298.
11. Baker PN, van der Meulen JH, Lewsey J, Gregg PJ. The role of pain and function in determining patient satisfaction after total knee replacement. Data from the National Joint Registry for England and Wales. The Journal of bone and joint surgery British volume 2007;89(7):893-900.
12. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010;468(1):57-63.
13. Noble PC, Conditt MA, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. Clin Orthop Relat Res 2006;452:35-43.
14. Nam D, Nunley RM, Barrack RL. Patient dissatisfaction following total knee replacement: a growing concern? Bone Joint J 2014;96-b(11 Supple A):96-100.
15. DeFrance M, Scuderi G. Are 20% of Patients Actually Dissatisfied Following Total Knee Arthroplasty? A Systematic Review of the Literature. J Arthroplasty 2022;
16. Gunaratne R, Pratt DN, Banda J, Fick DP, Khan RJK, Robertson BW. Patient Dissatisfaction Following Total Knee Arthroplasty: A Systematic Review of the Literature. J Arthroplasty 2017;32(12):3854-3860.
17. Hafkamp FJ, Gosens T, de Vries J, den Oudsten BL. Do dissatisfied patients have unrealistic expectations? A systematic review and best-evidence synthesis in knee and hip arthroplasty patients. EFORT Open Rev 2020;5(4):226-240.
18. Kahlenberg CA, Nwachukwu BU, McLawhorn AS, Cross MB, Cornell CN, Padgett DE. Patient Satisfaction After Total Knee Replacement: A Systematic Review. Hss j 2018;14(2):192-201.
19. Marsh J, Joshi I, Somerville L, Vasarhelyi E, Lanting B. Health care costs after total knee arthroplasty for satisfied and dissatisfied patients. Canadian journal of surgery Journal canadien de chirurgie 2022;65(5):E562-e566.

20. Coyle J, Williams B. Seeing the wood for the trees: defining the forgotten concept of patient dissatisfaction in the light of patient satisfaction research. Int J Health Care Qual Assur Inc Leadersh Health Serv 1999;12(6-7):i-ix.
21. Hudak PL, McKeever P, Wright JG. The metaphor of patients as customers: implications for measuring satisfaction. J Clin Epidemiol 2003;56(2):103-8.
22. Wylde V, Dieppe P, Hewlett S, Learmonth ID. Total knee replacement: is it really an effective procedure for all? Knee 2007;14(6):417-23.
23. Wright RJ, Sledge CB, Poss R, Ewald FC, Walsh ME, Lingard EA. Patient-reported outcome and survivorship after Kinemax total knee arthroplasty. J Bone Joint Surg Am 2004;86(11):2464-70.
24. Bayley KB, London MR, Grunkemeier GL, Lansky DJ. Measuring the success of treatment in patient terms. Med Care 1995;33(4 Suppl):As226-35.
25. Lieberman JR, Dorey F, Shekelle P, et al. Differences between patients' and physicians' evaluations of outcome after total hip arthroplasty. J Bone Joint Surg Am 1996;78(6):835-8.
26. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. Clin Orthop Relat Res 1989;(248):13-4.
27. Janse AJ, Gemke RJ, Uiterwaal CS, van der Tweel I, Kimpen JL, Sinnema G. Quality of life: patients and doctors don't always agree: a meta-analysis. J Clin Epidemiol 2004;57(7):653-61.
28. Noble PC, Fuller-Lafreniere S, Meftah M, Dwyer MK. Challenges in outcome measurement: discrepancies between patient and provider definitions of success. Clin Orthop Relat Res 2013;471(11):3437-45.
29. Eckhard L, Munir S, Wood D, et al. The ceiling effects of patient reported outcome measures for total knee arthroplasty. Orthopaedics & traumatology, surgery & research : OTSR 2021;107(3):102758.
30. Dawson J, Beard DJ, McKibbin H, Harris K, Jenkinson C, Price AJ. Development of a patient-reported outcome measure of activity and participation (the OKS-APQ) to supplement the Oxford knee score. Bone Joint J 2014;96-b(3):332-8.
31. Witjes S, van Geenen RC, Koenraadt KL, et al. Expectations of younger patients concerning activities after knee arthroplasty: are we asking the right questions? Qual Life Res 2017;26(2):403-417.
32. Price AJ, Longino D, Rees J, et al. Are pain and function better measures of outcome than revision rates after TKR in the younger patient? Knee 2010;17(3):196-9.
33. Terwee CB, Peipert JD, Chapman R, et al. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. Qual Life Res 2021;30(10):2729-2754.
34. de Vet HC, Terwee CB. The minimal detectable change should not replace the minimal important difference. J Clin Epidemiol 2010;63(7):804-8.
35. Tubach F, Dougados M, Falissard B, Baron G, Logeart I, Ravaud P. Feeling good rather than feeling better matters more to patients. Arthritis Rheum 2006;55(4):526-30.
36. Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. Annals of the rheumatic diseases 2005;64(1):34-7.

Chapter 2

Translation, cross-cultural adaptation, reliability and construct validity of the Dutch Oxford Knee Score – Activity and Participation Questionnaire

Malou E.M. te Molder
Johanna E. Vriezekolk
Menno R. Bénard
Petra J.C. Heesterbeek

BMC Musculoskeletal Disorders. 2021;22(1): 700

Abstract

Background

Patients undergoing total knee arthroplasty (TKA) tend to be younger and tend to receive TKA at an earlier stage compared to 20 years ago. The Oxford Knee Score – Activity and Participation (OKS-APQ) questionnaire evaluates higher levels of activity and participation, reflecting activity patterns of younger or more active people. The purpose of this study was to translate the OKS-APQ questionnaire into Dutch, and to evaluate its measurement properties in pre- and postoperative TKA patients.

Methods

The OKS-APQ was translated and adapted according to the forward–backward translation multi step approach and tested for clinimetric quality. Floor and ceiling effects, structural validity, construct validity, internal consistency and test–retest reliability were evaluated using COSMIN quality criteria. The OKS-APQ, the Oxford Knee Score (OKS), the Short Form-36 (SF-36), a Visual Analogue Scale (VAS) for pain and the Forgotten Joint Score (FJS) were assessed in 131 patients (72 preoperative and 59 postoperative TKA patients), and the OKS-APQ was administered twice in 50 patients (12 preoperative and 38 postoperative TKA patients), after an interval of minimal 2 weeks.

Results

Floor effects were observed in preoperative patients. Confirmatory factor analyses (CFA) indicated a good fit of a 1-factor model by the following indices: (Comparative Fit Index (CFI): 0.97, Tucker-Lewis Index (TLI): 0.96 and Standardized Root Mean Square Residual (SRMR): 0.03). Construct validity was supported as > 75% of the hypotheses were confirmed. Internal consistency (Cronbach α's from 0.81 to 0.95) was good in the pooled and separate pre- and postoperative samples and test–retest reliability (Intraclass Correlation Coefficients (ICCs) from 0.63 – 0.85) were good in postoperative patients and moderate in preoperative patients. The standard Error of Measurements (SEMs) ranged from 8.5 – 12.2 and the Smallest Detectable Changes in individuals (SDCind) ranged from 23.5 – 34.0 (on a scale from 0 to 100).

Conclusions

Preliminary findings suggest that the Dutch version of the OKS-APQ is reliable and valid for a Dutch postoperative TKA patient sample. However, in a preoperative TKA sample, the OKS-APQ seems less suitable, because of floor effects and lower test–retest reliability. The Dutch version of the OKS-APQ can be used alongside the OKS to discriminate among levels of activity and participation in postoperative patients.

Keywords

Oxford knee score – activity and participation questionnaire, Dutch version, Total knee arthroplasty, Patient-reported outcome measure, Translation, Validation

Introduction

The Oxford knee score (OKS) questionnaire is a validated patient-reported outcome measure (PROM), developed for patients undergoing Total Knee Arthroplasty (TKA).[1] The OKS was developed in 1998 to reflect patients' perception of knee pain and functional impairment after TKA.[1] Nowadays, patients undergoing TKA tend to be younger and receive TKA at an earlier stage compared to 20 years ago.[2,3] According to the latest annual report of the Dutch Arthroplasty Register (LROI), 17% of the primary TKAs were performed in patients younger than 60 years old.[4] Younger patients with an active lifestyle, have higher expectations of the outcome after the procedure.[5,6] Patients want to stay active and engaged in their social and recreational activities up to and after retiring.[3] Regaining a higher level of participation in social and recreational activities becomes more important for patients after TKA.[3] This implies that besides pain and disability, higher levels of activity and participation have become an important outcome domain. For that reason, the original OKS was extended with an additional one-dimensional scale, the Oxford Knee Score – Activity and Participation Questionnaire (OKS-APQ), to better monitor changes in activity and participation levels after TKA.[5]

While TKA procedures are highly successful because it is proven to relieve pain and to improve function, still a significant proportion of approximately 20% of the patients is not satisfied after surgery.[7,8] The fulfilment of preoperative patient expectations clearly seems to play an important role in patient satisfaction.[7] Especially in younger patients because they expect to perform better on many activities of daily life, work and leisure time after TKA.[6] Therefore, it is important to use questionnaires that reflect patients' perception of their quality of life, including activities relevant for younger patients.

Following the Dutch Orthopaedic Association (NOV) TKA guideline (2014), patients undergoing TKA in our hospital complete a standard set of questionnaires (e.g. OKS, KOOS-PS, NRS and EQ-5D) for routine outcome monitoring.[9] This set of questionnaires no longer seems sufficient due to concerns about existing ceiling effects of the OKS and EQ-5D in younger patients. Meaning that highest scores on the OKS and EQ-5D would not necessarily reflect treatment satisfaction in the younger patient group.[6] This has been observed by Dawson et al. and in response they developed the OKS-APQ to extend the OKS.[5] Before the OKS-APQ may be used in Dutch clinical practice for outcome monitoring or used for research purposes in young and active patients, the OKS-APQ needs to be translated into Dutch and the measurement properties need to be examined.

The unidimensional, eight-item OKS-APQ evaluates activity and participation levels (e.g. sports, dancing, and participation in activities with friends and family) that fit activity patterns of younger or more active patients. It consists of four highly valued activities and four items concerning performance and awareness (e.g. timing and adjustments of activities).[5] Besides the original English version of the OKS-APQ and a Chinese version[10], the questionnaire has not been translated and validated in other languages including Dutch. The original OKS-APQ has shown to be a valuable complement to the OKS, particularly where further detail regarding the levels of activity and participation are required.[5]

The present study aimed to translate the OKS-APQ into the Dutch language and to assess the unidimensionality of the instrument, the test-retest reliability, internal consistency, construct validity and floor and ceiling effects, in pre- and postoperative TKA patients.[11]

Methods

We performed a translation of the OKS-APQ into Dutch and prospectively evaluated the measurement properties of the Dutch version. Measurement properties were evaluated using COSMIN quality criteria.[11]

Procedure of translation

The OKS-APQ questionnaire was translated from English to Dutch according to the advised forward-backward translation multi step approach for translation as described by Beaton et al.[12,13] First, two independent native Dutch translators (DT1 and DT2) translated the OKS-APQ questionnaire to Dutch (forward translation). A definitive version (V12) was based on consensus within a team of translators, health professionals and researchers. Second, two native English translators (ET1 and ET2), blinded to the original English version by Dawson et al., independently re-translated the Dutch version (V12) into English (backwards translation).[5] Third, the definitive Dutch version of the OKS-APQ was made after a consensus meeting with the team. During the last step, the comprehensibility and interpretability of the definitive version was pilot-tested in a subset of 5 preoperative and 5 postoperative TKA patients. These patients completed the questionnaire at home and were asked to make notes if they thought a question was difficult to understand. Hereafter, a researcher contacted all 10 patients by telephone to discuss the difficulties and to ascertain the meaning that patients attributed to the OKS-APQ items.[14] Recruiting patients for the pilot test was stopped after 10 patients because no issues regarding the OKS-APQ items were reported or emerged. Therefore, no alterations were made to the instruction or questions.

Patients

As a rule of thumb, at least 100 patients were required and we aimed to include preoperative and postoperative patients. The preoperative study sample was recruited from the waiting list for TKA. Postoperative patients were selected from the outpatient registry. Inclusion criteria for the study participants were: clinically diagnosed with knee OA, age above 18 years, scheduled for TKA within the next 6 weeks or had undergone TKA between 6 and 12 months ago. Patients unable to speak Dutch and understand Dutch written language were excluded. All patients would undergo or underwent TKA at the department of orthopaedics at the Sint Maartenskliniek in Nijmegen. The study was assessed by the local hospital review committee. No ethical approval was sought for, as this study was not subject to the Dutch medical research involving human subjects act. All patients gave their written informed consent prior to study participation.

Questionnaires

Besides completing the OKS-APQ, patients completed additional condition-specific questionnaires commonly used in pre- and postoperative TKA patients for hypothesis testing purposes between January 2017 and December 2019. All preoperative patients completed the following four questionnaires: the OKS-APQ, the Oxford Knee Score (OKS)[15], the Short Form-36 (SF-36)[16], and a Visual Analogue Scale for pain[17]. Postoperative patients also completed an additional fifth questionnaire, the Forgotten Joint Score (FJS).[18] All patients were asked to complete the OKS-APQ questionnaire for a second time, after a minimum of two weeks, which was considered appropriate for the test-retest reliability.[11]

Oxford Knee Score - Activity and Participation (OKS-APQ)

The OKS-APQ eight-item questionnaire was developed to measure higher levels of activity and participation and is recommended to be used to complement the OKS as an additional scale.[5] Items are scored on a five-point Likert scale, ranging from 0 “strongly agree” to 4 “strongly disagree”. Total summary score ranges from 0 to 32, and scores are converted to a 0 to 100 measure.[5] A lower total sum score represents lower levels of activity and participation.[5]

Oxford Knee Score (OKS)

The OKS 12-item questionnaire has been developed for patients undergoing TKA to evaluate the patients’ perception of pain and functional impairment in the knee.[15] The Dutch questionnaire consists of 12 questions and it is possible to derive separate OKS pain and function subscales.[15] Responses are scored on a 5-point Likert scale, ranging from 0 “significant disability” to 4 “no problem”, in which the final score is an aggregate, sum score for pain and function.[19] The total scores ranges from 0 to 48; a lower OKS sum score represents poor function and more pain. The Dutch OKS has good measurement properties[15], however ceiling effects were demonstrated in postoperative patients.[20,21]

MOS Short Form 36 (SF-36v2)

The Dutch SF-36 version 2 is a 36-item questionnaire assessing health-related Quality of Life (QoL). It consists of eight dimensions that are aggregated to two summary scores: Physical Component Score (PCS) and Mental Component Score (MCS) (both 0-100).[16] The SF-36 is widely used and has shown to be reliable and valid in the Dutch general population.[16,22,23] A lower score represents a lower level of QoL.[15]

Visual analogue scale for pain (VAS pain)

The Dutch VAS for pain is a single item scale assessing the intensity of pain in the knee during the past 2 to 3 days. The 100-mm VAS is simple to use, and has already been applied in different populations and settings.[17] The score varies from 0 (no pain), to 100 (worst pain). It has shown to be valid and reliable.[17]

Forgotten Joint Score (FJS-12)

The Dutch 12-item Forgotten Joint Score (FJS-12) questionnaire evaluates the patients’ ‘joint awareness’ during activities of daily living (i.e. stair climbing, walking and gardening). The responses were scored on a five-point Likert scale, ranging from 0 “never” to 4 “mostly”. Item scores were summed and converted to a 0 to 100 scale, a low total sum score reflects that the patient is not able to forget the affected/replaced joint during activities of daily living.[18] The Dutch FJS-12 has shown to be a reliable and valid questionnaire.[24]

Methodological testing & statistical analysis

Kolmogorov Smirnov test was used to test the normality of the OKS-APQ items, OKS-APQ total score and other PROM total scores. Descriptive statistics were used to summarize the data; mean and standard deviation (SD) or median (25^{th} – 75^{th} percentile) for continuous variables

and counts and percentages for categorical variables, and to investigate the frequencies of missing data. All statistical analyses were performed in STATA version 13.0 (StataCorp, College Station, Texas). A p-value < 0.05 was considered statistically significant for all analyses.

Validity

Validity is the degree to which the Dutch OKS-APQ measures the construct(s) it purports to measure. To evaluate validity, floor and ceiling effects, structural validity and construct validity were measured.

Floor & ceiling effects

Another quality criterion is the absence of floor and ceiling effects. Presence of floor and ceiling effects on the OKS-APQ may influence the test-retest reliability, and construct validity of the questionnaire.[25] Patients with the lowest or highest possible score cannot be distinguished from each other, thus reliability is reduced.[25] Floor and ceiling effects, in pre- and postoperative samples separately, were determined by calculating the number of individuals that obtained the lowest (0) or highest (100) scores possible and were considered present if more than 15% of the patients achieved the highest or lowest total summary score.[25] In addition, floor and ceiling effects on item-level were determined to provide information about the item distribution.

Structural validity

Confirmatory factor analyses (CFA) was used to validate the 1-factor structure of the original English version of the OKS-APQ.[5] We examined the comparative fit index (CFI; values ranging from 0.90 to 0.95 indicate an adequate fit and values greater than or equal to 0.95 indicate a good fit), the Tucker-Lewis Index (TLI; values ranging from 0.90 to 0.95 indicate an adequate fit and values greater than or equal to 0.95 indicate a good fit), the root mean square error of approximation (RMSEA; values ranging from 0.05 to 0.08 represent adequate fit and values less than 0.05 indicate good fit) and the standardized root mean squared residual (SRMR; values less than or equal to 0.08 indicate good fit) to assess goodness of fit of this model. CFA was assessed using the pooled sample (pre- and postoperative patients).

Construct validity

Validity is the degree to which the OKS-APQ measures the construct it supposes to measure. Since there is no gold standard in the measurement of PROMs, validity was measured as construct validity.[25] Construct validity refers to the extent to which the OKS-APQ was related to other measures based on theoretically derived, predefined hypotheses. Construct validity was supported when at least 75% of the results are in accordance with the predefined hypotheses (Table 1).[25] Construct validity was expressed by assessing Pearson correlation coefficients or the nonparametric Spearman's correlation coefficients.[25] The strength of the correlations was interpreted as "weak" (r = 0.10 - 0.30), "moderate" (r = 0.31 - 0.50) or "strong" (r = 0.51 – 1.00).[26] Predefined hypotheses were formulated for the pooled and separate pre- and postoperative samples.

Table 1. Predefined hypotheses for evaluating the construct validity of the Dutch OKS-APQ

Hypothesis
A. Strong positive correlation (r > 0.50):
1. A strong positive correlation between OKS-APQ and OKS (pooled, pre- and postoperative patients);
2. A strong positive correlation between OKS-APQ and FJS (in postoperative patients);
3. A strong positive correlation between OKS-APQ and PCS (SF-36) (pooled, pre- and postoperative patients);
B. Moderate to strong negative correlation (r > 0.31):
4. A moderate to strong negative correlation between OKS-APQ and VAS pain (pooled, pre- and postoperative patients);
C. Weak to moderate positive correlation (r 0.10 – 0.50):
5. A weak to moderate positive correlation between OKS-APQ and MCS (SF-36) (pooled, pre- and postoperative patients);

Abbreviations: OKS-APQ, Oxford knee score – Activity & Participation Questionnaire; FJS, Forgotten joint score; OKS, Oxford knee score; SF-36, 36-Item Short Form Health Survey Questionnaire; VAS for pain, Visual Analogue Scale; PCS, Physical Component Score; MCS, Mental Component Score

Reliability

Reliability is the degree to which the Dutch OKS-APQ is free from measurement error. To evaluate reliability, internal consistency, test-retest reliability, the measurement error and the smallest detectable change were calculated.

Internal consistency

Internal consistency is a measure to evaluate to what extent the eight items of the Dutch OKS-APQ refer to the same underlying construct.[25] Internal consistency of the Dutch version of the OKS-APQ was determined by calculating the Cronbach's alpha.[25] A Cronbach's alpha between 0.7 and 0.9 for the eight items of the OKS-APQ indicates good internal consistency.[25] The Cronbach's alpha was measured on the pooled sample and the separate pre- and postoperative samples.

Test-retest reliability

Test-retest reliability involves the degree to which the results of the Dutch OKS-APQ are consistent across repeated measurements.[25] To evaluate the reliability of the Dutch OKS-APQ, we calculated intraclass correlation coefficients (ICCs) with a 95% confidence interval (95% CI). In addition, we provided the different variance components to show the systematic differences between the two timepoints in preoperative and postoperative patients. More specific, we used the ICC two-way random effects model type agreement to measure the reliability.[25] An ICC equal to and larger than 0.7 is generally accepted as good.[25] ICCs were calculated for the separate pre- and postoperative samples.

Measurement error & Smallest detectable change

The measurement error is the systematic and random error of a participant's score that is not attributed to true changes in the construct to be measured.[11] The standard error of measurement (SEM) was calculated using the square root of the error variance.[14,25]

The smallest detectable change (SDC) reflects the smallest individual change in score that can be interpreted as a real change in one individual (SDC_{ind}). This was calculated by the SEM * 1.96 * √2.[14,25] The SDC_{ind} can be divided by √n (n = sample size) to calculate the SDC in a group of patients (SDC_{group}).[14,25] SEM and SDC were calculated for the separate pre- and postoperative samples.

Results

Demographic data

A total of 131 patients were included, with mean age 66.3 (9.4) years, of whom 72 were preoperative patients with OA prior to TKA, and 59 were postoperative patients ≥6 months after TKA (Table 2). Both the pooled data and the separated pre- and postoperative samples were not normally distributed ($p < 0.05$). The missing values per item and for the total scores ranged from: 0 to 5.34% missing values, with the latter only for VAS pain. All missing items on the OKS-APQ (Table 3), OKS and SF-36 were imputed as recommended with patient-specific mean values of completed items. Because 10% missing data for a variable is considered acceptable[27], we performed the analyses without further evaluation or adjustment of the other variables.

Table 2. Patient Characteristics

Sociodemographic	Pooled sample (n = 131)	Preoperative sample (n = 72)	Postoperative sample (n = 59)
Age; mean (SD),(yr)	66.3 (9.4)	66.2 (9.3)	66.4 (9.6)
Self-report measures; median (25^{th} – 75^{th} percentile)			
OKS - Activity & Participation (OKS-APQ) (scale 0-100)	21.9 (6.3 – 56.3)	10.9 (0 – 23.4)	62.5 (25 – 84.4)
Oxford Knee Score (OKS) (scale 0-48)	29 (20 – 39)	22 (15 – 29)	39 (30 – 44)
VAS Pain (scale 0-100)	30 (10.5 – 63.5)	59 (31 – 74)	11 (4 – 28)
Quality of life			
Physical component (SF-36-PCS) (scale 0-100)	34.1 (27.8 – 40.8)	30.6 (25.8 – 34.8)	39.8 (33.6 – 46.8)
Mental component (SF-36-MCS) (scale 0-100)	52.8 (42.5 – 57.2)	50.6 (41.7 – 56.4)	53.7 (48.0 – 57.3)
Forgotten Joint Score (FJS) (scale 0-100)	NA	NA	37.5 (14.6 – 60.4)

Abbreviations: SD indicates standard deviation; OKS APQ Oxford Knee Score - Activity and Participation Questionnaire; OKS Oxford Knee Score; VAS Visual Analogue Scale; SF-36, 36-item Short Form Health Survey Questionnaire; SF-36-MCS, Mental Component Score; SF-36-PCS, Physical Component Score; FJS Forgotten Joint Score

Floor & ceiling effects

Floor effects were observed for the individual items and summary score of the OKS-APQ in the preoperative patient sample (Table 3). Twenty one patients (29.2%) scored the lowest level of activity and participation. No ceiling effect was observed for the summary score and individual items. In the postoperative patient sample, no floor and ceiling effects were observed for the summary score. On item level, both floor and ceiling effects were observed, however responses were much more evenly distributed (Table 3).

Structural validity

CFA indicated a good fit of a 1-factor model by the following indices: CFI: 0.97, TLI: 0.96, and SRMR: 0.03. However, RMSEA: 0.11, was greater than 0.08.

Construct validity

Construct validity was assessed with Spearman's rho correlations and showed a strong positive correlation for the OKS in both pre- and postoperative patients and a strong positive correlation for the FJS-12 and PCS of the SF-36 in postoperative patients (Table 4). The OKS-APQ showed a moderate to strong negative correlation for the VAS pain and a weak to moderate positive correlation for the MCS of the SF-36 in both pre- and postoperative patients.

Internal consistency

The item-total correlations were calculated for each item (Table 3). Internal consistency was appropriate; Cronbach alpha values exceeded 0.70 for the pooled and separate samples of pre- and postoperative patients (Table 5).

Test-retest reliability

Fifty patients (12 preoperative and 38 postoperative patients) completed the questionnaires for a second time, after a minimum of two weeks. The median scores (25^{th} – 75^{th} percentile) for the test and retest of the OKS-APQ, the ICCs and variance components are presented in Table 5. The OKS-APQ showed good test-retest reliability in the postoperative samples with an ICC of 0.85. The ICC in the preoperative sample was lower with smaller between-subject variability in preoperative patients (Table 5).

Measurement error & Smallest detectable change

SEM, SDC_{ind} and SDC_{group} in the pre- and postoperative patients are presented in Table 5.

Table 3. Characteristics of the Dutch OKS-APQ

Items	Item-Total Correlation (pooled sample)	Missing, n (%) (pooled sample)	Floor, n (%) (preoperative sample)	Ceiling, n (%) (preoperative sample)	Floor, n (%) (postoperative sample)	Ceiling, n (%) (postoperative sample)
It is a problem for me to do activities (e.g. sports, dancing, walking) to the level I want, because of my knee	0.84	0 (0.0%)	64 (89%)	1 (1.4%)	19 (32.2%)	7 (11.9%)
It is a problem for me to carry heavy things (e.g. items at work, shopping or a child), because of my knee	0.83	0 (0.0%)	41 (56.9%)	1 (1.4%)	12 (20.3%)	15 (25.4%)
I need to modify my work or everyday activities, because of my knee	0.92	4 (3.1%)	42 (60%)	3 (4.3%)	11 (19.3%)	18 (31.6%)
I need to plan carefully before going out for the day because of my knee (e.g. taking painkillers, using a knee brace or checking that there will be places to sit down)	0.89	0 (0.0%)	44 (61.1%)	2 (2.8%)	9 (15.3%)	27 (45.8%)
It is a problem for me to fully take part in activities with friends and family, because of my knee	0.85	0 (0.0%)	34 (47.2%)	1 (1.4%)	11 (18.6%)	18 (30.5%)
It is a problem for me to walk at the pace I would like, because of my knee	0.88	0 (0.0%)	62 (86.1%)	10 (13.9%)	16 (27.1%)	12 (20.3%)
It is a problem for me to twist or turn, as my knee may give way or be painful	0.87	0 (0.0%)	46 (63.9%)	3 (4.2%)	10 (17%)	23 (39%)
It is a problem for me that I need to take longer to do everyday activities, because of my knee	0.88	1 (0.8%)	29 (40.9%)	8 (11.3%)	7 (12.1%)	18 (31%)

Table 4. Construct validity of the Dutch OKS-APQ

Predefined Hypothesis	Spearman correlation* (*r*)		
	Pooled sample	Preoperative sample	Postoperative sample
1. A strong positive correlation between OKS-APQ and OKS;	0.83	0.63	0.80
2. A strong positive correlation between OKS-APQ and FJS;	NA	NA	0.74
3. A strong positive correlation between OKS-APQ and PCS (SF-36);	0.65	0.40	0.59
4. A moderate to strong negative correlation between OKS-APQ and VAS pain;	-0.68	-0.43	-0.63
5. A weak to moderate positive correlation between OKS-APQ and MCS (SF-36);	0.47	0.50	0.40

Abbreviations: OKS-APQ, Oxford Knee Score – Activity & Participation Questionnaire; OKS, Oxford Knee Score; FJS, Forgotten Joint Score; SF-36, 36-Item Short Form health survey questionnaire; VAS for pain, Visual Analogue Scale; PCS, Physical Component Score; MCS, Mental Component Score

*All correlations $P < 0.001$

Table 5. Reliability of the Dutch OKS-APQ

Study sample		Pooled sample (n = 131)	Preoperative sample (n = 72)	Postoperative sample (n = 59)
Cronbach α		0.95	0.81	0.95
Test-retest sample			**Preoperative sample (n = 12)**	**Postoperative sample (n = 38)**
OKS-APQ test *median* *(25th – 75th percentile)*			9.38 (0.00 – 29.69)	67.19 (27.34 – 82.03)
OKS-APQ retest *median* *(25th – 75th percentile)*			15.63 (0.78 – 28.13)	62.50 (21.88 – 78.13)
ICC (95% CI)			0.63 (0.10 – 0.88)	0.85 (0.72 – 0.92)
Variance components	σ^2_p		125.62	820.31
	σ^2_o		0.00	2.08
	$\sigma^2_{residual}$		72.02	147.49
SEM			8.49	12.23
SDC_{ind}			23.53	34.00
SDC_{group}			6.79	5.52

Abbreviations: ICC, Intraclass Correlation Coefficient; SEM, Standard Error of Measurement; SDC_{ind}, Smallest Detectable Change in one individual; SDC_{group}, Smallest Detectable Change in a group

σ^2_p: The variance of the patients (i.e. the systematic differences between the 'true' scores of the patients; σ^2_o: variance due to systematic differences between observers/timepoints; $\sigma^2_{residual}$: Residual variance (i.e. random error variance)

Discussion

In general, the Dutch OKS-APQ indicated to be an understandable, reliable and valid unidimensional PROM to assess activity and participation levels in post-operative TKA patients. No floor and ceiling effects were observed for the summary score of the OKS-APQ in postoperative patients. However, floor effects were observed in preoperative patients indicating that the Dutch OKS-APQ is not able to discriminate among the lowest levels of activity and participation in the preoperative situation solely based on the OKS-APQ. Furthermore, internal consistency was good in the pooled and separate samples. Test-retest reliability was good in the postoperative sample, however, was moderate in the preoperative sample. In the overall sample, structural validity indicated satisfactory 1-factor model fit.

Cross-cultural translation

The cross-cultural translation and adaptation procedure in this study yielded a clear, understandable Dutch version of the OKS-APQ. Content validity, including the relevance and comprehensiveness was not evaluated in this study. Likewise, content validity ratio (CVR) and content validity index (CVI) were not determined. Witjes et al., however, showed that the OKS-APQ was rated as an important and relevant questionnaire for younger Dutch TKA patients.[6] Since content validity is an important measurement property according to the recent COSMIN study design checklist for patient-reported outcome measurement instruments[28], further investigation of the Dutch OKS-APQ is advised to evaluate its content validity with patients and experts.

Floor & ceiling effects

In general, the patterns of observed floor and ceiling effects of the Dutch OKS-APQ for the summary score and at item level were consistent with the original OKS-APQ and the Chinese version of the OKS-APQ.[5,10] The floor effects found in the preoperative sample might be explained by the fact that these patients were awaiting a TKA and therefore report severe complaints/functional limitations. In the postoperative sample both floor and ceiling effects were present at item level, that can be explained by the varying rehabilitation course after TKA. Some of these patients were still rehabilitating after 6 months, while others were already fully recovered.

Reliability and structural validity

Confirmatory factor analysis of the Dutch OKS-APQ in our pooled sample of pre- and postoperative patients confirmed the unidimensional structure of the original OKS-APQ[5] as was reflected by good fit indices. Nevertheless, as we can not rule out bias by pooling the data of pre- and postoperative patients, it is important to replicate these findings in larger, separate pre- and postoperative samples. Furthermore, in line with the original and Chinese version of the OKS-APQ, the internal consistency of the Dutch OKS-APQ was good for the pooled and separate sample of pre- and postoperative patients. The test-retest reliability was good for the postoperative sample (ICC = 0.85). In contrast to the Chinese OKS-APQ validation findings that showed an excellent test-retest reliability (ICC = 0.94) in a sample of 30 preoperative patients, we found a moderate test-retest reliability in the preoperative sample (ICC = 0.63). The ICC in the preoperative patient sample was lower than the ICC of the postoperative patient sample which may be explained by the small preoperative sample size in our study, and in turn, the smaller between-subject variability in preoperative patients (see Table 5). Since the ICC is a relative measure depending on both the between-subject variability and test-retest variability similar test-retest variability in combination with smaller between-subject variability resulted in a lower ICC value. Since a sample size of at least 50 patients is recommended for examination of the test-retest reliability of a health measurement instrument[25], further investigation of the Dutch OKS-APQ in larger pre- and postoperative samples is recommended to firmly establish its test-retest reliability.

Construct validity

The construct validity of the Dutch OKS-APQ was confirmed as more than 75% of our hypotheses were supported. In line with other research[5,10], the Dutch OKS-APQ strongly correlated with knee specific questionnaires (e.g. OKS and AKSS), and the general SF-36 physical component score. Overall, this suggests that the OKS-APQ, OKS, AKSS and the physical component score of the SF-36 measure similar constructs.

Clinical implications

For clinical practice, this study shows that the Dutch OKS-APQ is able to discriminate among postoperative patients whereas ceiling effects were previously found for the OKS in postoperative patients.[20,21] The developers of the OKS-APQ recommend to use the OKS-APQ to complement the OKS as an additional scale.[5] Caution in interpretation of preoperative OKS-APQ evaluation is warranted because of floor effects found in preoperative patients. Evidently, preoperative scores are needed to evaluate effects of surgical interventions as TKA. In addition, the OKS-APQ may provide support for transferring patients to transmural care (e.g. physiotherapy or social work) when patients are still not satisfied with the prosthesis because of problems in social participation and recreational activities including sports. This may be subject for future investigations.

Limitations

Limited by our cross-sectional study design and small group sample sizes, several measurement properties of the Dutch OKS-APQ could not be evaluated. Further validation studies in larger samples are recommended to more extensively evaluate the content validity (e.g. exploring the relevance and comprehensiveness with patients and experts), structural validity of the OKS-APQ (e.g. testing structure equivalence of the Dutch OKS-APQ in pre- and postoperative TKA patients separately), the reliability and precision of the OKS-APQ (e.g. test-retest reliability in larger pre- and postoperative samples and differential item functioning using item response modelling), responsiveness (e.g. testing the validity of change scores of the Dutch OKS-APQ), interpretability (e.g. by relating the SDC and SEM to the minimal important change (MIC)) and predictive validity. Furthermore, our findings were based on a sample of patients who were treated in a specialized hospital, this should be taken into account when generalizing to other samples or settings.

Conclusion

Preliminary findings suggest that the Dutch version of the OKS-APQ is reliable and valid for a Dutch postoperative TKA patient sample. However, in a preoperative TKA sample, the OKS-APQ seems less suitable, because of floor effects and lower test-retest reliability. The Dutch version of the OKS-APQ can be used alongside the OKS to discriminate among levels of activity and participation in postoperative patients.

Acknowledgements

THE OXFORD KNEE SCORE – ACTIVITY AND PARTICIPATION QUESTIONNAIRE (OKS-APQ)

References

1. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. The Journal of bone and joint surgery British volume 1998;80(1):63-9.
2. Kurtz SM, Lau E, Ong K, Zhao K, Kelly M, Bozic KJ. Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030. Clin Orthop Relat Res 2009;467(10):2606-12.
3. Witjes S, Gouttebarge V, Kuijer PP, van Geenen RC, Poolman RW, Kerkhoffs GM. Return to Sports and Physical Activity After Total and Unicondylar Knee Arthroplasty: A Systematic Review and Meta-Analysis. Sports Med 2016;46(2):269-92.
4. (LROI) DAR. Retrieved from: https://www.lroi-report.nl/knee/patient-characteristics/. Date last accessed: 3 May 2021.
5. Dawson J, Beard DJ, McKibbin H, Harris K, Jenkinson C, Price AJ. Development of a patient-reported outcome measure of activity and participation (the OKS-APQ) to supplement the Oxford knee score. Bone Joint J 2014;96-b(3):332-8.
6. Witjes S, van Geenen RC, Koenraadt KL, et al. Expectations of younger patients concerning activities after knee arthroplasty: are we asking the right questions? Qual Life Res 2017;26(2):403-417.
7. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010;468(1):57-63.
8. Van Onsem S, Van Der Straeten C, Arnout N, Deprez P, Van Damme G, Victor J. A New Prediction Model for Patient Satisfaction After Total Knee Arthroplasty. J Arthroplasty 2016;31(12):2660-2667.e1.
9. (NOV) DOA. TKA guideline. 2014. Retrieved from: http://www.richtlijnendatabase.nl/totale_knieprothese_2014/totale_knieprothese_-_korte_beschrijving_2014.html. Date last accessed: 3 May 2021.
10. Chen C, Wang W, Wu H, et al. Cross-cultural translation and validation of the Chinese Oxford Knee Score and the Activity and Participation Questionnaire. Journal of orthopaedic surgery (Hong Kong) 2020;28(2):2309499020910668.
11. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19(4):539-49.
12. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000;25(24):3186-91.
13. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46(12):1417-32.
14. de Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Cambridge University Press; 2011.
15. Haverkamp D, Breugem SJ, Sierevelt IN, Blankevoort L, van Dijk CN. Translation and validation of the Dutch version of the Oxford 12-item knee questionnaire for knee arthroplasty. Acta orthopaedica 2005;76(3):347-52.
16. Aaronson NK, Muller M, Cohen PD, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. J Clin Epidemiol 1998;51(11):1055-68.
17. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis care & research 2011;63 Suppl 11:S240-52.
18. Hamilton DF, Loth FL, Giesinger JM, et al. Validation of the English language Forgotten Joint Score-12 as an outcome measure for total hip and knee arthroplasty in a British population. Bone Joint J 2017;99-b(2):218-224.

19. Murray DW, Fitzpatrick R, Rogers K, et al. The use of the Oxford hip and knee scores. The Journal of bone and joint surgery British volume 2007;89(8):1010-4.
20. Harris K, Dawson J, Gibbons E, et al. Systematic review of measurement properties of patient-reported outcome measures used in patients undergoing hip and knee arthroplasty. Patient Relat Outcome Meas 2016;7:101-8.
21. Reito A, Järvistö A, Jämsen E, et al. Translation and validation of the 12-item Oxford knee score for use in Finland. BMC Musculoskelet Disord 2017;18(1):74.
22. Anderson RT, Aaronson NK, Wilkin D. Critical review of the international assessments of health-related quality of life. Qual Life Res 1993;2(6):369-95.
23. VanderZee KI, Sanderman R, Heyink J. A comparison of two multidimensional measures of health status: the Nottingham Health Profile and the RAND 36-Item Health Survey 1.0. Qual Life Res 1996;5(1):165-74.
24. Shadid MB, Vinken NS, Marting LN, Wolterbeek N. The Dutch version of the Forgotten Joint Score: test-retesting reliability and validation. Acta Orthop Belg 2016;82(1):112-8.
25. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60(1):34-42.
26. Cohen J. A power primer. Psychol Bull 1992;112(1):155-9.
27. Bennett DA. How can I deal with missing data in my study? Aust N Z J Public Health 2001;25(5):464-9.
28. Mokkink LB, Prinsen CA, Patrick DA, Alonso J, Bouter LM, de Vet HC. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. Retrieved from: www.cosmin.nl. Date last accessed: 15 Jun 2021.

Chapter 3

Definitions of poor outcome after total arthroplasty: An inventory review

Malou E.M. te Molder
José M.H. Smolders
Petra J.C. Heesterbeek
Cornelia H.M. van den Ende

BMC Musculoskeletal Disorders. 2020;21(1): 378

Abstract

Background

A significant proportion of patients experiences poor response (i.e. no or little improvement) after total knee arthroplasty (TKA) because of osteoarthritis. It is difficult to quantify the proportion of patients who experiences poor response to TKA, as different definitions of, and perspectives (clinician's and patient's) on poor response are being used. The aim of this study was therefore to review the literature and summarize definitions of poor response to TKA.

Methods

A systematic search was performed to identify and review studies that included dichotomous definitions of poor outcome after primary TKA. The type, amount and combination of domains (e.g. functioning), outcome measures, type of thresholds (absolute/relative, change/cut-off), values and moments of follow-up used in definitions were summarized.

Results

A total of 47 different dichotomous definitions of poor response to TKA were extracted from 2163 initially identified studies. Thirty-six definitions incorporated one domain, seven definitions comprised two domains and four definitions comprised three domains. Eight different domains were used in identified definitions: pain, function, physical functioning, quality of life (QoL), patient satisfaction, anxiety, depression and patient global assessment. The absolute cut-off value was the most common type of threshold, with large variety in value and timing of follow-up.

Conclusions

Our inventory review shows that definitions of poor response to TKA are heterogeneous. Our findings stresses the need for an unambiguous definition of poor response to draw conclusions about the prevalence of poor-responders to TKA across hospitals and countries, and to identify patients at risk.

Keywords

Total knee arthroplasty, Treatment outcome, Poor outcome, Patient-reported outcome measures, Osteoarthritis knee

Introduction

Total knee arthroplasty (TKA) is considered to be a cost-effective intervention for the treatment of advanced knee osteoarthritis (OA). Nonetheless, increasing evidence shows that a significant proportion of patients experiences a poor response to TKA (i.e. show no or little improvement) in terms of chronic knee pain[1,2], functional disability[3,4], poor quality of life (QoL)[5], and dissatisfaction after TKA.[6-13] However, both in research and in clinical practice it is a challenge to identify those patients with an unfavourable course after TKA, as different definitions of non-response are being used.

The comparison of research findings on the effectiveness of TKA across studies and countries is hampered by the fact that different concepts for (poor) response for outcome after TKA are being used.[14] Various concepts or definitions of (poor) response are developed on group level, using mean changes, to describe improvement over time in patient cohorts. However, these concepts are not useful to measure (poor) response on an individual level. Furthermore, there might be differences in perspective of concepts of (poor) response among different stakeholders (i.e. physician, patient, clinical researcher or health insurer). Physicians usually focus on aspects of a dimension based on their clinical evaluation (e.g. stability, range of motion and alignment), while patients focus on the functionality of the knee during daily life activity.[8] Moreover, the view of physicians and patients on the desired magnitude of improvement after TKA are not always consistent[6], as poor correlations were found between physician-assessed and patient-reported outcome measures (PROMs).[15] Research findings on the outcome of TKA are predominantly based on single continuous outcome measures assessing one construct, without taking all potential relevant constructs into account. A preliminary set of domains for total joint replacement (TJR) clinical trials was proposed by international experts that included pain, function, patient satisfaction, revision, adverse events, and death[16], but no propositions were made to what extent these domains should be incorporated in a definition of (poor) response to TKA.

Currently, it is unclear which definitions of poor response to primary TKA are used in the literature. This insight can help to reach consensus on an unambiguous definitions of poor response. The need to use a combined endpoint incorporating relevant constructs, and a relevant amount of change based on multiple, clinical outcome perspectives (including physicians' and patients' outcome perspective) to accurately describe poor-response to TKA[17], has been recognized in the literature.

The aim of this study was to review and summarize dichotomous definitions of poor response as dichotomous definitions allow to make inferences about the prevalence of poor outcome and comparisons of TKR outcome across hospitals, countries and over time.

We expect a variety of definitions and outcomes used to define poor response after primary TKA. Therefore this study systematically map definitions of poor response to primary TKA in the literature.

Methods

Search Strategy

Previously used definitions for poor outcome after TKA in the literature, from 2000 up to October 2019 were identified and reviewed. We followed the following strategy to systematically identify definitions of poor outcome:

- An electronic search strategy was performed to retrieve systematic reviews (SRs) until 2016 in the PubMed Database, EMBASE and PsycINFO, on the outcome of TKA using search terms as "knee, arthroplasty, replacement, prosthesis, outcome measures, score and failure". A detailed search strategy can be found in Figure 1 and an example of the search strings is available (Supplementary file 1).
- To include more recent publications, articles from 2016 or later, were searched for definitions of poor outcome of TKA using the same search terms as for the SRs.
- Duplicates were excluded and search results were screened on title and abstract.
- All reference lists of included SRs were hand searched for relevant articles.
- Studies were assessed for eligibility by the in- and exclusion criteria.
- Subsequently, full texts screening of the eligible studies was carried out, for definitions of poor outcome after primary TKA.

Study selection

Two reviewers (CvdE, MtM) independently selected eligible SRs and eligible recent publications from 2016 or later, describing a dichotomous definition for poor outcome after TKA in a certain domain or combination of domains. The reviewers (CvdE, MtM) independently selected eligible studies from the references lists of the, based on their titles, SRs and thereafter the eligibility criteria for studies.

Eligibility criteria studies

Selected studies were assessed by the following eligibility criteria: 1) TKA patient population diagnosed with osteoarthritis; 2) primary TKA; 3) incorporating a dichotomous definition of poor outcome after primary TKA; 4) utilizing a controlled or prospective observational design (cohort and registry studies).

There were no restrictions on 1) sample size; 2) the type of primary implant; 3) follow-up moment of outcome; and 4) studies with missing data regarding primary- and secondary outcomes.

Studies with other TKA procedures (i.e. revision TKA, uni- or bicompartimental replacements) as study intervention, studies not written in the English language and articles without an abstract and/or access to the full text manuscript, were excluded.

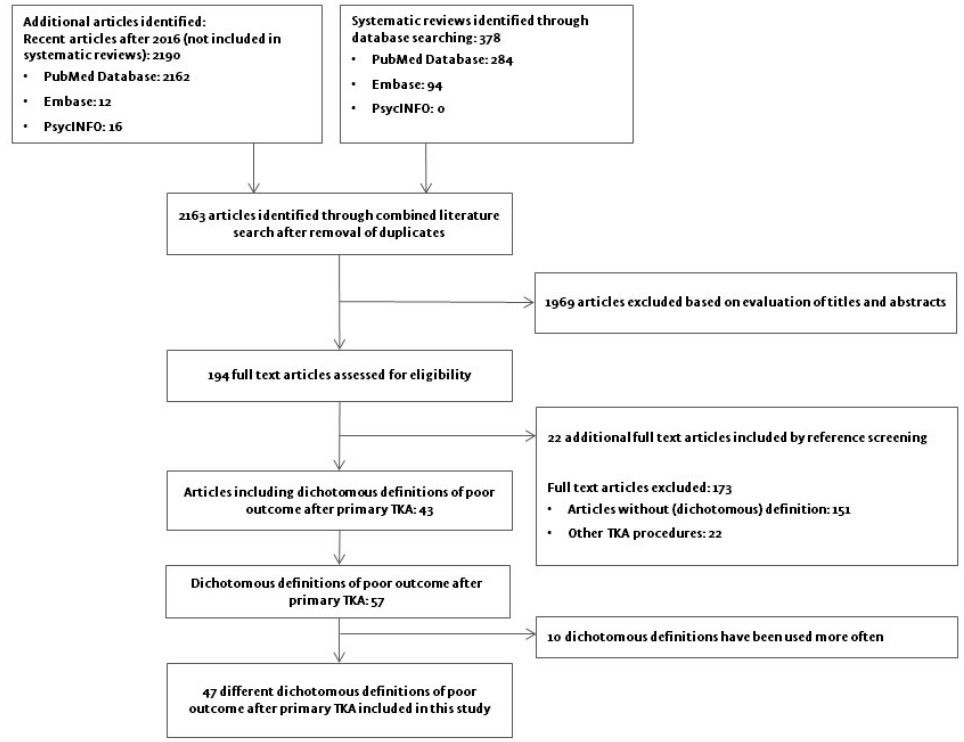


Figure 1. Flowchart of the study and definition selection

Selection of definitions

Full text evaluation of the eligible studies was carried out by one reviewer (MtM), to select eligible dichotomous definitions. When eligible studies used definitions from other articles, data on the definition was extracted from the original article. Criteria for definitions to be included in the review were: 1) included a threshold for dichotomizing; and 2) related to the concept of poor, worse or non-responder outcome as stated by the authors of that publication. Any uncertainty about the final selection of definitions and the extraction of data on the definitions was discussed with the second reviewer (CvdE) to achieve consensus.

Data collection and synthesis

We extracted all dichotomous definitions of poor outcome after TKA using a data extraction form. Definitions were grouped according to outcome domains (e.g. pain, function, patient satisfaction), the following data were extracted:

- Study characteristics: author names, date of publication and length of study follow-up.
- Characteristics of the definition of poor outcome after primary TKA: number of domains, type of domains, outcome measure(s) used, type of threshold (absolute/relative cut-off value/change), value and time points.
- Additional background information on selected definitions of poor response provided by authors.

Results

After removal of duplicates, a total of 329 SRs and 1834 articles after 2016 were screened on title and abstract (Fig. 1). The full texts of 48 SRs and 146 articles after 2016 were assessed for eligibility. In total 43 articles included 57 definitions of poor outcome after primary TKA. Forty-seven different dichotomous definitions of poor outcome after primary TKA were included in this study (Supplementary file 2).

There were 36 different unidimensional definitions, seven different definitions included two domains and four different definitions incorporated three domains. Eight different domains were identified in the 47 different definitions: pain, function, physical functioning, QoL, patient satisfaction, anxiety, depression, and patient global assessment. Pain (*N* = 28 different definitions), physical functioning (*N* = 17 different definitions) and patient satisfaction (*N* = 13 different definitions) domains were most frequently used in identified definitions of poor outcome after TKA. Patient satisfaction was only measured as unidimensional definitions with wide variation in wording of questions and answering categories. An absolute cut-off value of a certain outcome measure (*N* = 42 different definitions) was the most common type of threshold.

In the 47 different dichotomous definitions of poor outcome after primary TKA, we identified 14 different single item questions, two different self-composite question, one physician-assessed outcome measure, eight different PROMs and five different mixed outcome measures (combination of physician-assessed and patient reported). Single item questions measure one construct by asking for example the following question: 'How satisfied are you with the outcome after TKA?' Self-composite questions refer to composite questions of multiple items: 'Overall limitations was defined as moderate/severe, if a patient had ≥ activities (walking, stairs, rising chair) with moderate or severe limitations (reference, <2 limitations)'. The Western Ontario and McMaster University Osteoarthritis Index (WOMAC) (*N* = 11 different dichotomous definitions) and the Visual Analogue Scale (VAS) (*N* = 5 different dichotomous definitions) were the most frequently used measurement instruments after the single item questions. In 27 of the 57 selected definitions, additional background information was provided on the choice for the definitions and/or thresholds being used, other definitions were not substantiated (Supplementary file 3).

Discussion

To our knowledge, this is the first inventory review that summarizes definitions of poor outcome after primary TKA presented in the literature. We found a total of 47 different definitions varying in nature and number of outcome domains involved, the type of response and the magnitude of change. A total of eight different dimensions were used in identified definitions of poor outcome: pain, function, physical functioning, health-related quality of life (HRQOL), patient satisfaction, anxiety, depression and patient global assessment. Patient satisfaction was used as single domain with a wide variation in wording of questions and answering categories. The absolute cut-off value was the most common type of threshold, with large variety in value and timing of follow-up. Our review stresses the need for an unambiguous, dichotomous definition for poor response after TKA to enable comparisons of the effectiveness of TKA among studies and among countries.

A remarkable finding of our review was that the majority of definitions used to describe poor outcome incorporate only one or two outcome domains. This finding does not seem to correspond with the conclusions by the OMERACT-OARSI initiative and the International Consortium for Health Outcomes Measurement (ICHOM). The OMERACT proposed a simplified set of responder criteria for (non-surgical) treatment of OA in clinical pharmaceutical trials. This set of responder criteria comprises relative and absolute changes in three domains: pain, physical functioning and patient global assessment.[18,19] Also, the OMERACT TJR Working Group proposed a set of core domains (pain, function, patient satisfaction, revision, adverse events, and death) to evaluate joint replacement in randomized controlled trials.[20] Parallel, ICHOM has developed a set of patient-centered outcome measures and case-mix factors for evaluating, comparing and improving the treatment (both surgical and non-surgical) of patients with hip and knee OA, focusing on outcomes that matter to patients.[21] Pain, function, HRQOL and work status formed the core outcome domains, after a modified Delphi process.[21] Corresponding to this standard set and these responder criteria it seems important to measure poor response to TKA within multiple constructs to cover important key outcome domains to patients.

A great variety of thresholds is being used to measure poor response, ranging from an absolute cut-off point regarding patient dissatisfaction to composite measures incorporating relative changes or a MCID less than a certain value. Some studies used the inverse of the OMERACT-OARSI responder criteria "nonresponse" as a definition of poor outcome.[22-26] However, it is questionable whether "poor response" is the true opposite of "clinically meaningful response" as this definition implies that patients with smaller improvements will be part of the poor response group. The study by Mahler et al. showed a clear asymmetric magnitude of change, with a lower amount of change for patients who reported being worsened compared to the amount of improvement in patients who reported being improved.[27] In our opinion the amount of absolute or relative change in relevant constructs is therefore an important aspect of definitions of poor outcome.

In our opinion, strict, dichotomous definitions are necessary to interpret data on group level and to compare TKR outcome among hospitals, countries and over time. However, dichotomous data implies reduction of data and is therefore, less suitable for identifying factors underlying poor outcome. In particular, for individual patients, continuous outcomes are more suitable to monitor and evaluate specific health outcomes.

Patient dissatisfaction

Patient satisfaction was used as single domain with a wide variation in wording of questions and answering categories, most frequently measured by single item questions (non-validated instruments).[14] However, patient satisfaction is a multidimensional construct that may represent either satisfaction with outcome (e.g. knee function) of TKA or the process of care delivery, which all can be influenced by patients' expectations.[14,28] Halawi et al. explored subjective reasons for patient dissatisfaction after TJR and found different causes of patient dissatisfaction. The most common causes for dissatisfaction after TKA were persistent pain, functional limitation, surgical complication and reoperation, staff or quality of care issues and unmet expectations.[28] It is likely that different factors influence the construct of patient satisfaction, and therefore it is important to determine the different determinants that contribute to patient satisfaction after TKA according to the perspective of patients and orthopaedic surgeons.

Different perspectives

The variety in definitions of poor outcome used could reflect different perspectives of physician, patient and clinical researcher. There are many studies reporting on the disagreement between the patient and physician in terms of their satisfaction with surgery.[7,29,30] It is conceivable that physicians tend to focus on aspects of their clinical evaluation (e.g. stability, range of motion and alignment), while patients are more likely to focus on the functionality of the knee during daily life activity. Moreover, the view of physicians and patients on the desired magnitude of improvement after TKA is not always concordant, as poor correlations were found between physician-assessed and PROMs.[6] Furthermore, most outcome measures have been developed according to the medical research perspective, which mainly address knee-specific measures like pain and function scores, and scarcely address mental functioning and consequences for social participation.[31] So far, the choice for definitions to describe response or non-response after TKA has been dominated by non-comprehensive physician-based scoring systems and PROMs in quantitative research but the perspectives of patients and orthopaedic surgeons regarding the definition of poor response have been relatively neglected.

Additional background information

In 27 of the 57 selected definitions additional background information was provided to justify the choice for the definition and/or thresholds being used. Background information was extracted from the original publications. In particular, definitions of patient dissatisfaction were not substantiated and arbitrarily dichotomized. This study has some limitations, as our searches for relevant articles were systematic but the data extraction was performed by a single reviewer. Although any uncertainty about the selection of definitions and the extraction of data on the definitions was discussed with the second reviewer (CvdE). This inventory review does provide a complete overview of definitions of poor response after TKA that could be of interest to a large group of physicians and researchers involved in defining outcomes after TKA. Furthermore, only studies published in English language were included. For this reason, it cannot be ruled out that some studies were not identified (language bias).

In conclusion, this inventory review shows that many different heterogeneous definitions, incorporating several domains, for poor response to primary TKA are being used in the literature. Future research should focus on the perspectives and perceptions of orthopaedic surgeons and patients about constructs underlying poor response to TKA. Our findings stress the need for an consensus-based unambiguous, dichotomous definition of poor response to draw conclusions about the prevalence of poor-responders to TKA across hospitals and countries, and to identify patients at risk.

References

1. Lavand'homme P, Thienpont E. Pain after total knee arthroplasty: a narrative review focusing on the stratification of patients at risk for persistent pain. Bone Joint J 2015;97-b(10 Suppl A):45-8.
2. Wylde V, Dennis J, Gooberman-Hill R, Beswick AD. Effectiveness of postdischarge interventions for reducing the severity of chronic pain after total knee replacement: systematic review of randomised controlled trials. BMJ open 2018;8(2):e020368.
3. Luttjeboer JS, Bénard MR, Defoort KC, van Hellemondt GG, Wymenga AB. Revision Total Knee Arthroplasty for Instability-Outcome for Different Types of Instability and Implants. J Arthroplasty 2016;31(12):2672-2676.
4. Rodriguez-Merchan EC. Instability following total knee arthroplasty. Hss j 2011;7(3):273-8.
5. Canovas F, Dagneaux L. Quality of life after total knee arthroplasty. Orthopaedics & traumatology, surgery & research : OTSR 2018;104(1s):S41-s46.
6. Baker PN, van der Meulen JH, Lewsey J, Gregg PJ. The role of pain and function in determining patient satisfaction after total knee replacement. Data from the National Joint Registry for England and Wales. The Journal of bone and joint surgery British volume 2007;89(7):893-900.
7. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010;468(1):57-63.
8. Bullens PH, van Loon CJ, de Waal Malefijt MC, Laan RF, Veth RP. Patient satisfaction after total knee arthroplasty: a comparison between subjective and objective outcome assessments. J Arthroplasty 2001;16(6):740-7.
9. Kauppila AM, Kyllönen E, Ohtonen P, Leppilahti J, Sintonen H, Arokoski JP. Outcomes of primary total knee arthroplasty: the impact of patient-relevant factors on self-reported function and quality of life. Disabil Rehabil 2011;33(17-18):1659-67.
10. Lindberg MF, Miaskowski C, RustøEn T, Rosseland LA, Cooper BA, Lerdal A. Factors that can predict pain with walking, 12 months after total knee arthroplasty. Acta orthopaedica 2016;87(6):600-606.
11. Scott CE, Howie CR, MacDonald D, Biant LC. Predicting dissatisfaction following total knee replacement: a prospective study of 1217 patients. The Journal of bone and joint surgery British volume 2010;92(9):1253-8.
12. Van Onsem S, Van Der Straeten C, Arnout N, Deprez P, Van Damme G, Victor J. A New Prediction Model for Patient Satisfaction After Total Knee Arthroplasty. J Arthroplasty 2016;31(12):2660-2667.e1.
13. Van Onsem S, Verstraete M, Dhont S, Zwaenepoel B, Van Der Straeten C, Victor J. Improved walking distance and range of motion predict patient satisfaction after TKA. Knee Surg Sports Traumatol Arthrosc 2018;26(11):3272-3279.
14. Kahlenberg CA, Nwachukwu BU, McLawhorn AS, Cross MB, Cornell CN, Padgett DE. Patient Satisfaction After Total Knee Replacement: A Systematic Review. Hss j 2018;14(2):192-201.
15. Kane RL, Saleh KJ, Wilt TJ, Bershadsky B. The functional outcomes of total knee arthroplasty. J Bone Joint Surg Am 2005;87(8):1719-24.
16. Singh JA, Dohm M, Sprowson AP, et al. Outcome Domains and Measures in Total Joint Replacement Clinical Trials: Can We Harmonize Them? An OMERACT Collaborative Initiative. J Rheumatol 2015;42(12):2496-502.
17. Price AJ, Longino D, Rees J, et al. Are pain and function better measures of outcome than revision rates after TKR in the younger patient? Knee 2010;17(3):196-9.
18. Pham T, van der Heijde D, Altman RD, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. Osteoarthritis Cartilage 2004;12(5):389-99.
19. Smith TO, Mansfield M, Hawker GA, et al. Uptake of the OMERACT-OARSI Hip and Knee Osteoarthritis Core Outcome Set: Review of Randomized Controlled Trials from 1997 to 2017. J Rheumatol 2019;46(8):976-980.

20. Singh JA, Dowsey MM, Dohm M, et al. Achieving Consensus on Total Joint Replacement Trial Outcome Reporting Using the OMERACT Filter: Endorsement of the Final Core Domain Set for Total Hip and Total Knee Replacement Trials for Endstage Arthritis. J Rheumatol 2017;44(11):1723-1726.
21. Rolfson O, Wissig S, van Maasakkers L, et al. Defining an International Standard Set of Outcome Measures for Patients With Hip or Knee Osteoarthritis: Consensus of the International Consortium for Health Outcomes Measurement Hip and Knee Osteoarthritis Working Group. Arthritis care & research 2016;68(11):1631-1639.
22. Dowsey MM, Robertsson O, Sundberg M, Lohmander LS, Choong PFM, A WD. Variations in pain and function before and after total knee arthroplasty: a comparison between Swedish and Australian cohorts. Osteoarthritis Cartilage 2017;25(6):885-891.
23. Dowsey MM, Spelman T, Choong PF. Development of a Prognostic Nomogram for Predicting the Probability of Nonresponse to Total Knee Arthroplasty 1 Year After Surgery. J Arthroplasty 2016;31(8):1654-60.
24. Escobar A, Gonzalez M, Quintana JM, et al. Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. Osteoarthritis Cartilage 2012;20(2):87-92.
25. Riddle DL, Golladay GJ, Jiranek WA, Perera RA. External Validation of a Prognostic Model for Predicting Nonresponse Following Knee Arthroplasty. J Arthroplasty 2017;32(4):1153-1158.e1.
26. Weber M, Craiovan B, Woerner ML, Schwarz T, Grifka J, Renkawitz TF. Predictors of Outcome After Primary Total Joint Replacement. J Arthroplasty 2018;33(2):431-435.
27. Mahler E, den Broeder AA, Woodworth TG, et al. How should worsening in osteoarthritis be defined? Development and initial validation of preliminary criteria for clinical worsening in knee and hip osteoarthritis. Scand J Rheumatol 2017;46(5):396-406.
28. Halawi MJ, Jongbloed W, Baron S, Savoy L, Williams VJ, Cote MP. Patient Dissatisfaction After Primary Total Joint Arthroplasty: The Patient Perspective. J Arthroplasty 2019;34(6):1093-1096.
29. Harris IA, Harris AM, Naylor JM, Adie S, Mittal R, Dao AT. Discordance between patient and surgeon satisfaction after total joint arthroplasty. J Arthroplasty 2013;28(5):722-7.
30. Lieberman JR, Dorey F, Shekelle P, et al. Differences between patients' and physicians' evaluations of outcome after total hip arthroplasty. J Bone Joint Surg Am 1996;78(6):835-8.
31. Bay S, Kuster L, McLean N, Byrnes M, Kuster MS. A systematic review of psychological interventions in total hip and knee arthroplasty. BMC Musculoskelet Disord 2018;19(1):201.

Supplementary material

Supplementary file 1. Example search terms – for PubMed

PUBMED SRs:

((((((((("Arthroplasty, Knee"[Mesh] OR total knee arthroplasty [title]))) AND ("Outcome Assessment (Health Care)"[Mesh] OR Outcome [tiab] OR result [tiab] OR "Patient Reported Outcome Measures"[Mesh] OR patient reported [tiab] OR outcome score [tiab] OR PROM [tiab] OR clinical [tiab] OR function* [tiab] OR "Patient Satisfaction"[Mesh] OR satisfaction [tiab] OR "Prosthesis Failure"[Mesh] OR failure [tiab] OR "Pain"[Mesh] OR pain [tiab] OR "Postoperative Complications"[Mesh] OR complication* [tiab] OR "Postoperative Period"[Mesh] OR post-operative [tiab] OR "Long Term Adverse Effects"[Mesh] OR long-term [tiab]))) AND systematic [sb])

Filter used: Text availability: Full text

PUBMED articles after 2016:

((((((((("Arthroplasty, Knee"[Mesh] OR total knee arthroplasty [title]))) AND ("Outcome Assessment (Health Care)"[Mesh] OR Outcome [tiab] OR result [tiab] OR "Patient Reported Outcome Measures"[Mesh] OR patient reported [tiab] OR outcome score [tiab] OR PROM [tiab] OR clinical [tiab] OR function* [tiab] OR "Patient Satisfaction"[Mesh] OR satisfaction [tiab] OR "Prosthesis Failure"[Mesh] OR failure [tiab] OR "Pain"[Mesh] OR pain [tiab] OR "Postoperative Complications"[Mesh] OR complication* [tiab] OR "Postoperative Period"[Mesh] OR post-operative [tiab] OR "Long Term Adverse Effects"[Mesh] OR long-term [tiab]))))

Filters used: Publication dates (From 2016/01/01 to 2019/05/14); Text availability: Full text

Supplementary file 2. Definitions of poor outcome after primary TKA

Domain	Outcome measures used	Absolute/relative cut-off value/ change	Value	Time points used
Unidimensional (single domain used)				
Pain	VAS pain scale (0-5) VAS pain scale (0-100)	*absolute cut-off*	• >0 *(Forsythe et al., 2008* (1)*)* • >40 *(Brander et al., 2003* (2)*; Brander et al., 2007* (3)*)*	24M 1, 3, 6, 12M
	VAS pain at rest scale (0-100)	*absolute cut-off*	≥1 *(Lundblad et al., 2008* (4)*)*	18M
	VAS pain with movement scale (0-100)	*absolute cut-off*	≥1 *(Lundblad et al., 2008* (4)*)*	18M
	WOMAC pain subscale	*relative change*	<50% improvement from baseline to 6M FU *(Riddle et al., 2010* (5)*)*	6M
	WOMAC pain subscale	*absolute change*	MCID ≤4 *(Riddle et al., 2010* (5)*)*	6M
	WOMAC pain subscale (0-100)	*absolute cut-off*	<100 and present for at least the past 3M *(Wylde et al., 2011* (6)*)*	3M
	WOMAC pain subscale	*absolute cut-off*	Moderate or worse response in any of the five questions of the pain scale *(Czurda et al., 2010* (7)*)*	at least 18M
	KSS pain subscale	*absolute cut-off*	≤30 *(Fisher et al., 2007* (8)*)*	12M
	IKSS pain subscale (0-50)	*absolute cut-off*	<30 *(Dowsey et al., 2012* (9)*)*	12, 24M
	McGill Pain Index (0-5)	*absolute cut-off*	>0 *(Forsythe et al., 2008* (1)*)*	24M
	BPI pain scale (0-10)	*absolute cut-off*	≥3 *(Masselin- Dubois et al., 2013* (10)*)*	3M
	AKSS pain subscale (0-7)	*absolute cut-off*	≥5 *(Elson et al., 2006* (11)*)*	5Y
	NRS pain scale (0-10)	*absolute cut-off*	>3 *(Pinto et al., 2013* (12)*)*	4 to 6M
	Single item question pain	*absolute cut-off*	Category: moderate and severe pain *(Singh et al., 2010/06* (13)*; Singh et al., 2014* (14)*)*	2, 5Y
Function (impairment)	Knee flexion function	*absolute cut-off*	<90° *(Fisher et al., 2007* (8)*; Boonen et al., 2016* (15)*)*	1, 2Y
	Single item question function	*absolute cut-off*	Categories: 'somewhat better, same, worse' *(Singh et al., 2010/06* (13)*)*	2, 5Y
Physical Functioning	IKSS functioning subscale (0-100)	*absolute cut-off*	<60 *(Dowsey et al., 2012* (9)*)*	12, 24M
	WOMAC functioning subscale	*relative change*	<50% improvement from baseline to 6M FU *(Riddle et al., 2010* (5)*)*	6M
	WOMAC functioning subscale	*absolute change*	MCID ≤15 *(Riddle et al., 2010* (5)*)*	6M
	Single item question functioning	*absolute cut-off*	Maximum walk time ≤15 min *(Yong-Hao Pua et al., 2016* (16)*)*	6M
	Self-composite question functioning	*absolute cut-off*	≥2 activities (walking, stairs, rising chair) with moderate-severe limitations *(Singh et al., 2010/04* (17)*; Singh et al., 2010/06* (13)*; Singh et al., 2014* (14)*)*	2, 5Y
	Self-composite question functioning	*absolute cut-off*	No *(Filbay et al.,2019* (18)*)*	12M
Satisfaction	VAS satisfaction scale (0-100)	*absolute cut-off*	≥50 *(Judge et al., 2012* (19)*)*	6M
	New KSS satisfaction subscale (0-40)	*absolute cut-off*	<20 *(Onsem et al., 2016* (20)*)*	3M
	Single item question satisfaction	*absolute cut-off*	• Dissatisfied *(Aggarwal et al., 2013* (21)*)* • Dissatisfied (=somewhat dissatisfied and very dissatisfied) *(Gandhi et al., 2009* (22)*; Filbay et al., 2019* (18)*)* • Dissatisfied (=unsure and dissatisfied) *(Scott et al., 2010* (23)*)* • Unsatisfied (=unsatisfied and very unsatisfied) *(Bierke et al., 2017* (24)*)* • Less satisfied (=moderately satisfied, neutral, moderately dissatisfied and very dissatisfied) *(Vissers et al., 2010* (25)*)* • Unsatisfied (=somewhat dissatisfied and dissatisfied) *(Escobar et al., 2011* (26)*)* • Dissatisfied (=uncertain and dissatisfied) *(Scott et al., 2016* (23)*)* • Dissatisfied (=no and I'm not sure) *(Jacobs et al., 2014* (27)*)* • Dissatisfied (=no) *(Kunze et al., 2018* (28)*)* • Not Satisfied *(Halawi et al., 2019* (29)*)*	12M 12M 12M 12M 6M 12M 12M 2 to 5 Y 12M 12M
	Single item question satisfaction	*relative cut-off*	≤ 50% *(Merle-Vincent et al., 2011* (30)*)*	24M
Multidimensional (two domains used)				

			≥ 8 *(Vissers et al., 2010* (25))	6M
			• Equivalent or higher scores than 5 for partial scores and 10 for the total score *(Caracciolo et al., 2005* (31))	Discharge
Anxiety + Depression	HADS anxiety & depression subscales (0-100)/ (partial scale: 0-18 & total scale: 0-36)	*absolute cut-off*		
Pain + Physical functioning	OKS pain & functioning subscales (0-48)	*absolute cut-off*	< 27 *(Seah et al., 2017* (32))	6M
	OKS pain & functioning subscales	*absolute change*	• MCID ≤ 5 *(Alzahrani et al., 2011* (33))	12M
	WOMAC pain & functioning subscales (0-100)	*absolute cut-off*	• MCID ≤ 6 *(Filbay et al.,2019* (18)) < 60 *(Katz et al., 2007* (34))	2Y
	WOMAC pain & functioning subscales	*relative change*	< 10% difference between the mean pre- and post-operative scores of the pain and function dimensions *(Nunez et al., 2007* (35); *Nunez et al., 2009* (36))	36M, 7Y
Multidimensional (three domains used)				
Pain + Function (impairment) + Global assessment	OMERACT-OARSI responder criteria (WOMAC pain & functioning subscales & global score)	*absolute & relative change*	Non-responder: < 50% improvement and less than an absolute improvement of 20 points in either pain or function OR if there was improvement in 2 of the 3 following: pain of < 20% and an absolute change of < 10, function < 20% and an absolute change of < 10, *global improvement of* < 20% and an absolute change of < 10 *(Dowsey et al., 2016* (37); *Riddle et al., 2017* (38); *Dowsey et al., 2017* (39); *Weber et al., 2018* (40))	12M
Pain + Function (impairment) + Physical functioning	WOMAC pain, stiffness & functioning subscales (0-100)	*absolute cut-off*	> 40.4 *(Lungu et al., 2014* (41))	6M
	WOMAC pain, stiffness & functioning subscales	*absolute change*	MCID < 7.5 *(Alzahrani et al., 2011* (33))	12M
Function (impairment) + Physical functioning + HR-QoL	KSS knee & function subscales & UCLA functioning scale & SF-12 physical & mental subscales	*absolute change*	Deterioration of ≥ 21 compared to preoperative *(Kokubun et al., 2017* (42))	6W, 12W, 6M

Abbreviations: AKSS, American Knee Society Score; BPI, Brief Pain Inventory; HADS, Hospital Anxiety and Depression Scale; HR-QoL, health related quality of life; IKSS, International Knee Society Score; KSS, Knee Society Score; New KSS, New Knee Society Score; NRS, Numerical Rating Scale; OKS, Oxford Knee Score; OMERACT-OARSI, Outcome Measures in Rheumatology committee and Osteoarthritis Research Society International committee; PRI, Pain Rating Index; SF-12, 12 Item Short Form Survey; UCLA, University of California-Los Angeles; VAS, Visual Analogue Scale; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index

Time points: W, weeks; M, months; Y, years

References

1. Forsythe M, Dunbar M, Hennigar A, Sullivan M. Prospective relation between catastrophizing and residual pain. Pain Res Manag. 2008;13(4):335–41.
2. Brander V, Stulberg S, Adams A, Harden R, Bruehl S, Stanos S, et al. Predicting Total Knee Replacement Pain. Clin Orthop Relat Res. 2003;(416):27–36.
3. Brander V, Gondek S, Martin E, Stulberg S. Pain and depression influence outcome 5 years after knee replacement surgery. Clin Orthop Relat Res. 2007;(464):21–6.
4. Lundblad H, Kreicbergs A, Jansson K. Prediction of persistent pain after total knee replacement for osteoarthritis. J Bone Joint Surg Br. 2008;90-B(2):166–71.
5. Riddle D, Wade J, Jiranek W, Kong X. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. Clin Orthop Relat Res. 2010;468(3):798–806.
6. Wylde V, Hewlett S, Learmonth I, Dieppe P. Persistent pain after joint replacement: Prevalence, sensory qualities, and postoperative determinants. Pain. 2011;152(3):566–72.
7. Czurda T, Fennema P, Baumgartner M, Ritschl P. The association between component malalignment and post-operative pain following navigation-assisted total knee arthroplasty: results of a cohort/nested case-control study. Knee Surg Sport Traumatol Arthrosc. 2010;18(7):863–9.
8. Fisher D, Dierckman B, Watts M, Davis K. Looks Good But Feels Bad: Factors That Contribute to Poor Results After Total Knee Arthroplasty. J Arthroplasty. 2007;22(6):39–42.
9. Dowsey M, Nikpour M, Dieppe P, Choong P. Associations between pre-operative radiographic changes and outcomes after total knee joint replacement for osteoarthritis. Osteoarthr Cartil. 2012;20(10):1095–102.
10. Masselin-Dubois A, Attal N, Fletcher D, Jayr C, Albi A, Fermanian J, et al. Are psychological predictors of chronic postsurgical pain dependent on the surgical model? A comparison of total knee arthroplasty and breast surgery for cancer. J Pain. 2013;14(8):854–64.
11. Elson D, Brenkel I. Predicting Pain After Total Knee Arthroplasty. J Arthroplasty. 2006;21(7):1047–53.
12. Pinto P, McIntyre T, Ferrero R, Almeida A, Araújo-Soares V. Risk Factors for Moderate and Severe Persistent Pain in Patients Undergoing Total Knee and Hip Arthroplasty: A Prospective Predictive Study. PLoS One. 2013;8(9):1–11.
13. Singh J, O'Byrne M, Colligan R, Lewallen D. Pessimistic explanatory style: a psychological risk factor for poor pain and functional outcomes two years after knee replacement. J Bone Joint Surg Br. 2010;92-B(6):799–806.
14. Singh J, Lewallen D. Are outcomes after total knee arthroplasty worsening over time? A time-trends study of activity limitation and pain outcomes. BMC Musculoskelet Disord. 2014;15(1):1–9.
15. Boonen B, Schotanus M, Kerens B, van der Weegen W, Hoekstra H, Kort N. No difference in clinical outcome between patient-matched positioning guides and conventional instrumented total knee arthroplasty two years post-operatively. Bone Jt J. 2016;98B(7):939–44.
16. Pua YH, Seah F, Clark R, Poon C, Tan J, Chong H. Development of a prediction model to estimate the risk of walking limitations in patients with total knee arthroplasty. J Rheumatol. 2016;43(2):419–26.
17. Singh J, O'Byrne M, Harmsen S, Lewallen D. Predictors of moderate-severe functional limitations after primary total arthroplasty (TKA): 4,701 TKAs at 2-years and 2,935 at 5-years. Osteoarthr Cartil. 2010;18(4):2–15.
18. Filbay S, Hons B. Evaluating patients' expectations from a novel patient-centered perspective predicts knee arthroplasty outcome. J Arthroplasty. 2019;33(7):2146–52.
19. Judge A, Arden N, Cooper C, Kassim javaid M, Carr A, Field R, et al. Predictors of outcomes of total knee replacement surgery. Rheumatology. 2012;51(10):1804–13.
20. van Onsem S, van der Straeten C, Arnout N, Deprez P, van Damme G, Victor J. A new prediction model for patient satisfaction after total knee arthroplasty. J Arthroplasty. 2016;31(12):2660–7.

21. Aggarwal A, Agrawal A. Mobile vs fixed-bearing total knee arthroplasty performed by a single surgeon. A 4- to 6.5-year randomized, prospective, controlled, double-blinded study. J Arthroplasty. 2013;28(10):1712–6.
22. Gandhi R, Davey J, Mahomed N. Patient Expectations Predict Greater Pain Relief with Joint Arthroplasty. J Arthroplasty. 2009;24(5):716–21.
23. Scott CEH, Howie CR, Macdonald D, Biant LC. Predicting dissatisfaction following total knee replacement: a prospective study of 1217 patients. J Bone Jt Surg Br. 2010;92(9):1253–8.
24. Bierke S, Petersen W. Influence of anxiety and pain catastrophizing on the course of pain within the first year after uncomplicated total knee replacement: a prospective study. Arch Orthop Trauma Surg. 2017;137(12):1735–42.
25. Vissers M, de Groot I, Reijman M, Bussmann J, Stam H, Verhaar J. Functional capacity and actual daily activity do not contribute to patient satisfaction after total knee arthroplasty. BMC Musculoskelet Disord. 2010;11(121):1–8.
26. Escobar A, Gonzalez M, Quintana J, Vrotsou K, Bilbao A, Herrera-Espiñeira C, et al. Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. Osteoarthr Cartil. 2012;20(2):87–92.
27. Jacobs C, Christensen C, Karthikeyan T. Patient and Intraoperative Factors Influencing Satisfaction Two to Five Years After Primary Total Knee Arthroplasty. J Arthroplasty. 2014;29(8):1576–9.
28. Kunze K, Akram F, Fuller B, Zabawa L, Sporer S, Levine B. Internal Validation of a Predictive Model for Satisfaction After Primary Total Knee Arthroplasty. J Arthroplasty. 2019;34(4):663–70.
29. Halawi M, Jongbloed W, Baron S, Savoy L, Williams V, Cote M. Patient dissatisfaction after primary total joint arthroplasty: the patient perspective. J Arthroplasty. 2019;34(6):1093–6.
30. Merle-Vincent F, Couris C, Schott A, Conrozier T, Piperno M, Mathieu P, et al. Factors predicting patient satisfaction 2 years after total knee arthroplasty for osteoarthritis. Jt Bone Spine. 2011;78(4):383–6.
31. Caracciolo B, Giaquinto S. Self-perceived distress and self-perceived functional recovery after recent total hip and knee arthroplasty. Arch Gerontol Geriatr. 2005;41(2):177–81.
32. Seah R, Lim W, Lo N, Yew A, Chong H, Yeo S. Unexplained Pain Post Total Knee Arthroplasty With an Oxford Knee Score ≥20 at 6 Months Predicts Good 2-Year Outcome. J Arthroplasty. 2017;32(3):807–10.
33. Alzahrani K, Gandhi R, DeBeer J, Petruccelli D, Mahomed N. Prevalence of clinically significant improvement following total knee replacement. J Rheumatol. 2011;38(4):753–9.
34. Katz J, Mahomed N, Baron J, Barrett J, Fossel A, Creel A, et al. Association of hospital and surgeon procedure volume with patient-centered outcomes of total knee replacement in a population-based cohort of patients age 65 years and older. Arthritis Rheum. 2007;56(2):568–74.
35. Núñez M, Núñez E, Luis del Val J, Ortega R, Segur J, Hernández M, et al. Health-related quality of life in patients with osteoarthritis after total knee replacement: Factors influencing outcomes at 36 months of follow-up. Osteoarthr Cartil. 2007;15(9):1001–7.
36. Núñez M, Lozano L, Núñez E, Segur J, Sastre S, Maculé F, et al. Total knee replacement and health-related quality of life: Factors influencing long-term outcomes. Arthritis Care Res. 2009;61(8):1062–9.
37. Dowsey M, Spelman T, Choong P. Development of a Prognostic Nomogram for Predicting the Probability of Nonresponse to Total Knee Arthroplasty 1 Year After Surgery. J Arthroplasty. 2016;31(8):1654–60.
38. Riddle D, Golladay G, Jiranek W, Perera R. External validation of a prognostic model for predicting nonresponse following knee arthroplasty. J Arthroplasty. 2017;32(4):1153–8.
39. Dowsey M, Robertsson O, Sundberg M, Lohmander L, Choong P, W-Dahl A. Variations in pain and function before and after total knee arthroplasty: a comparison between Swedish and Australian cohorts. Osteoarthr Cartil. 2017;25(6):885–91.
40. Weber M, Craiovan B, Woerner M, Schwarz T, Grifka J, Renkawitz T. Predictors of outcome after primary total joint replacement. J Arthroplasty. 2018;33(2):431–5.
41. Lungu E, Desmeules F, Dionne C, Belzile É, Vendittoli P. Prediction of poor outcomes six months following total knee arthroplasty in patients awaiting surgery. BMC Musculoskelet Disord. 2014;15(1):1–11.
42. Kokubun B, Manista G, Courtney P, Kearns S, Levine B. Intra-Articular Knee Injections Before Total Knee Arthroplasty: Outcomes and Complication Rates. J Arthroplasty. 2017;32(6):1798–802.

Supplementary file 3. Additional background information about selected definitions of poor response provided by authors

Definition	Additional information
VAS pain >0 (Forsythe, 2008 (33))	"Sample was dichotomized into definitive 'pain'(>0) and 'no pain (0)' groups, rather than arbitrarily dichotomizing the population as having mild or moderate pain."
VAS pain >40 (Brander, 2007 (35))	"Cut-offs were defined according to Dolan and Sutton(1)."
VAS pain at rest ≥1 (Lundblad, 2008 (36))	"Dichotomization was performed at the median value."
VAS pain with movement ≥1 (Lundblad, 2008 (36))	"Dichotomization was performed at the median value."
WOMAC pain <50% (Riddle, 2010 (32))	"Patients were dichotomized into those who did and those who did not improve by 50% or greater based on changes in WOMAC pain scores from baseline to 6 months post-surgery. Our rationale for this approach is that large sample studies generally indicate average improvement for patients is approximately 50% relative to initial scores."
WOMAC pain MCID ≤4 (Riddle, 2010 (32))	"For 6-month change scores, we dichotomized our sample based on whether changes exceed the 6-month minimal clinically important difference (MCID)(2). These changes were greater than 4 points for WOMAC pain."
IKSS pain <30 (Dowsey et al, 2012 (40))	"Pain was dichotomized into two categories based on severity at 12 months and 2 years (those with IKSS pain score ≥30) were classified as having none to mild pain and those with IKSS pain score <30 were classified as having moderate to severe pain(3)."
McGill Pain Index >0 (Forsythe et al, 2008 (33))	"Sample was dichotomized into definitive 'pain' and 'no pain' groups, rather than arbitrarily dichotomizing the population as having mild or moderate pain."
Brief Pain Inventory ≥3 (Masselin-Dubois, 2013 (41))	"Clinically meaningful pain was considered to be present if patients rated their average pain or pain right now as ≥3/10 on the BPI. This cutoff corresponds to at least moderate pain with a potential impact on physical or emotional functioning."
NRS pain >3 (Pinto, 2013 (43))	"Patients reporting significant "worst pain" levels in the surgical area (NRS >3) were considered as being PPSP positive, similar to previous studies (4). ..... this cut-off was based on previous recommendations considering the differential impact of pain levels above 3 (5,6)."
Single item question function: somewhat better, same, worse (Singh, 2010/06 (44))	"Patients in the category of 'much better' were compared with those in the reference category comprising 'somewhat better, same, worse', Based on the fact that, typically, TKR is an extremely successful procedure and most patients aim to achieve much better knee function than pre-operatively."
IKSS functioning <60 (Dowsey, 2012 (40))	According to Asif et al. (7).
WOMAC functioning <50% improvement from baseline to 6M FU (Riddle, 2010 (32))	"Patients were dichotomized into those who did and those who did not improve by 50% or greater based on changes in WOMAC function scores from baseline to 6 months postsurgery. Our rationale for this approach is that large sample studies generally indicate average improvement for patients is approximately 50% relative to initial scores."
WOMAC functioning MCID ≤15 (Riddle, 2010 (32))	"For 6-month change scores, we dichotomized our sample based on whether changes exceed the 6-month minimal clinically important difference (MCID)(2). These changes were greater than 15 points for WOMAC function."
New KSS, satisfaction <20 (Onsem, 2016 (51))	"A score of more than 20 meant that the patient selected "satisfied" at least one time and a score lower than 20 that the patient selected "dissatisfied" at least once. We considered 20 as the cut off value for satisfaction."
HADS anxiety & depression ≥8 (Vissers, 2010 (55))	According to Zigmond et al. (8).
HADS Equivalent or higher scores than 5 for partial scores and 10 for the total score (Caracciolo, 2005 (60))	"The cut-off scores for the Italian version are fixed at 5 for the partial scores and at 10 for the total score. Equivalent or higher scores classify subjects as psychologically distressed. The evaluations were established at admission and at discharge."
OKS pain & functioning <27 (Seah, 2017 (61))	"Unexplained pain after primary TKA and an Oxford knee score (OKS) less than 27 (which is considered a poor outcome (9) at 6 months."
OKS pain & functioning MCID≤5 (Alzahrani, 2011 (62))	"We defined the MCID for the OKS as a 5-point change or less, as suggested by Murray et al. (10)."
OKS pain & functioning MCID≤6 (Filbay, 2019 (49))	"The MIC for the OKS following TKA at an individual level, has been estimated to be 7-points (considered as the minimal amount of change necessary to distinguish between patients 'a little better' from those 'about the same' in a UK sample of 94,502 individuals undergoing knee arthroplasty) (11). To enhance interpretation of findings and maintain consistency with other binary outcomes, the OKS was dichotomized using a cut-off of 7-points, whereby patients reporting an improvement of 6-points or less on the OKS between baseline and one-year follow-up were categorized as 'not achieving OKS MIC."
WOMAC pain & functioning <60 (Katz, 2007 (63))	"Scores <60 indicated poor outcome. Sixty is a typical preoperative score (12)."

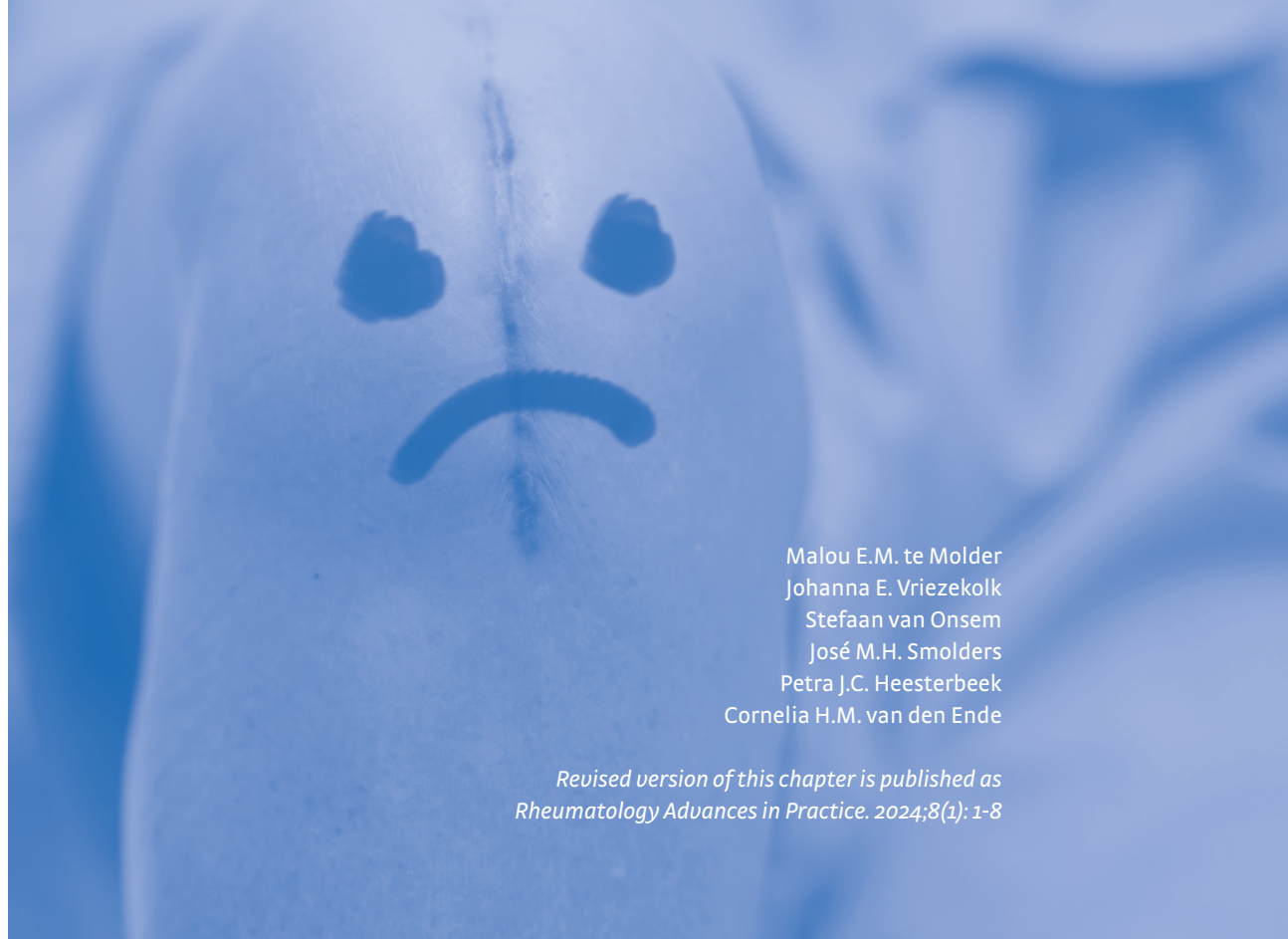
OMERACT-OARSI responder criteria (WOMAC, pain & functioning and global score): Non-responder: <50% improvement and less than an absolute improvement of 20 points in either pain or function OR if there was improvement in 2 of the 3 following: pain of <20% and an absolute change of <10, function <20% and an absolute change of <10, *global improvement of* <20% and an absolute change of <10 (Dowsey, 2016 (66) & 2017 (24); Riddle, 2017 (22); Weber, 2018 (26))	References made to Pham et al. (13).
WOMAC pain, stiffness & functioning >40.4 (Lungu, 2014 (67))	"As there is no universal agreement on what is considered poor outcome following TKA surgery, it was defined as the last quintile of the six-month postoperative WOMAC score (i.e. WOMAC score > 40.4); a satisfactory outcome was defined..... distribution (i.e. score ≤ 40.4)."
WOMAC pain, stiffness & functioning MCID<7.5 (Alzahrani, 2011 (62))	According to Mahomed et al. (14).

References

1. Dolan P, Sutton M. Mapping visual analogue scale health state valuations onto standard gamble and time trade-off values. Soc Sci Med. 1997;44(10):1519–30.
2. Escobar A, Quintana J, Bilbao A, Aróstegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. Osteoarthr Cartil. 2007;15(3):273–80.
3. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. Clin Orthop Relat Res. 1989;248(November (248)):13–4.
4. Vuorenmaa M, Ylinen J, Kiviranta I, Intke A, Kautiainen H, Mälkiä E, et al. Changes in pain and physical function during waiting time and 3 months after knee joint arthroplasty. J Rehabil Med. 2008;40(7):570–5.
5. Dihle A, Helseth S, Paul S, Miaskowski C. The exploration of the establishment of cutpoints to categorize the severity of acute postoperative pain. Clin J Pain. 2006;22(7):617–24.
6. Bodian C, Freedman G, Hossain S, Eisenkraft J, Beilin Y. The visual analog scale for pain. Clinical significance in postoperative patients. Anesthesiology. 2001;95(6):1356–61.
7. Asif S, Choon D. Midterm results of cemented Press Fit Condylar Sigma total knee arthroplasty system. J Orthop Surg (Hong Kong). 2005;13(3):280–4.
8. Zigmond A, Snaith R. The Hospital Anxiety and Depression Scale. Acta psychiatr scand. 1983;67:361–70.
9. Pynsent P, Adams D, Disney S. The Oxford hip and knee outcome questionnaires for arthroplasty. J Bone Jt Surg Br. 2005;87-B(2):241–8.
10. Murray D, Fitzpatrick R, Rogers K, Pandit H, Beard D, Carr A, et al. The use of the Oxford hip and knee scores. J Bone Jt Surg Br. 2007;89-B(8):1010–4.
11. Beard D, Harris K, Dawson J, Doll H, Murray D, Carr A, et al. Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. J Clin Epidemiol. 2015;68(1):73–9.
12. Fortin P, Clarke A, Joseph L, Liang M, Tanzer M, Ferland D, et al. Outcomes of total hip and knee replacement: Preoperative functional status predicts outcomes at six months after surgery. Arthritis Rheum. 1999;42(8):1722–8.
13. Pham T, Van Der Heijde D, Lassere M, Altman R, Anderson J, Bellamy N, et al. Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria. J Rheumatol. 2003;30(7):1648–54.
14. Mahomed N, Davis A, Hawker G, Badley E, Davey J, Syed K, et al. Inpatient compared with home-based rehabilitation following primary unilateral total hip or knee replacement: A randomized controlled trial. J Bone Jt Surg Am. 2008;90(8):1673–80.

Chapter 4

Exploration of adverse consequences of TKA by patients and knee specialists: A qualitative study



Malou E.M. te Molder
Johanna E. Vriezekolk
Stefaan van Onsem
José M.H. Smolders
Petra J.C. Heesterbeek
Cornelia H.M. van den Ende

Revised version of this chapter is published as Rheumatology Advances in Practice. 2024;8(1): 1-8

Abstract

Objectives

A successful outcome according to the knee specialist is not a guarantee for treatment success as perceived by patients. This study explored outcome expectations and experiences of patients with osteoarthritis (OA) before and after total knee arthroplasty (TKA) surgery and knee specialists that may contribute to the negative appraisal of its effect, and differences in views between patients and knee specialists.

Methods

Semi-structured interviews were held in Belgium and the Netherlands. Twenty-five patients (2 without TKA indications, 11 on the waiting list for TKA and 12 postoperative TKA), and fifteen knee specialists (9 orthopaedic surgeons, 1 physician assistant, 1 nurse practitioner, and 4 physiotherapists) were interviewed. Conversations were audio recorded, transcribed verbatim, and analysed using thematic analysis following the grounded theory approach. Separate analyses were conducted for patients and knee specialists.

Results

Patients were focused on the arduous process of getting used to the prosthesis, lingering pain, awareness of the artificial knee and limitations they experience during valued and daily activities, while knee specialists put emphasis on surgical failure, unexplained pain, limited walking ability and impairments that limit patients' physical functioning.

Conclusion

This study provides a comprehensive overview of potential adverse consequences from the perspective of both patients and knee specialists. Improving patients' awareness and expectations of adaptation to the knee prosthesis needs to be considered.

Keywords

Total knee arthroplasty, Poor outcome, Qualitative study

Introduction

Total knee arthroplasty (TKA) is considered to be a successful and cost-effective intervention for treatment of advanced symptomatic osteoarthritis (OA).[1] However, despite improvement in knee pain and disability, 15-20% of the patients report being dissatisfied with their prosthetic knee due to insufficient pain relief, loss of function and limitations in physical functioning.[2-5] To improve these dissatisfaction rate, a clear and valid definition of what poor outcome after TKA entails is needed. Currently, various dichotomous definitions comprising one or more outcome dimensions are used to quantify the proportion of patients with a poor response to TKA.[6]

Importantly, the current various definitions for poor outcome reflect the perspective of physicians and researchers. Previous research shows that a successful outcome according to the physician is not a guarantee for treatment success as perceived by patients.[3,7,8] A definition of poor response that is supported by patients and physicians is crucial to allow benchmarking across (inter)national institutions for quality improvement and will facilitate improved (shared) decision making.

Recently, in a qualitative study using nominal group technique, patients identified refractory pain after total joint arthroplasty (TJA) as more important than surgical failure (i.e. complications, revisions).[9] However, group responses/themes about failure were ranked and prioritized by 42 postoperative patients from only one high-volume centre. The latter study did not provide insight in differences between the views of physicians and patients about TKA failure. Hence, the purpose of our multicentre, qualitative study was to explore outcome expectations and experiences of patients with OA and knee specialists (i.e. orthopaedic surgeons, physician assistants (PA), nurse practitioners (NP) and physiotherapists) after TKA surgery that may contribute to the negative appraisal of its effect. Secondary aim was to explore whether these views differ between patients and knee specialists.

Methods

Study design and setting

A cross-sectional, multicentre qualitative study was performed using semi-structured interviews. To support our objective of exploring outcome expectations and experiences on adverse consequences of TKA, methods of a constructivist grounded theory approach[10] with thematic analysis[11] were applied. The interviews were held in patients with OA and health care providers with expertise on knee replacement surgery and its rehabilitation. Patients were recruited from one Belgian and two Dutch hospitals; knee specialists were recruited from various hospitals and physiotherapy practices in Belgium and the Netherlands. The Standards for Reporting of Qualitative Research (SRQR) checklist[12] was used to ensure complete and transparent reporting (Supplementary Table S1).

Participants

Interviews with patients

Purposive sampling was used to capture three different patient subgroups 1) knee OA patients without surgical indication, 2) patients scheduled for TKA, 3) patients 1 to 5 years after TKA. Patients in subgroup 1 were included because they may have different outcome expectations regarding a TKA procedure compared to patients scheduled for TKA. Regarding the latter subgroup, purposive sampling was directed towards achieving different outcome experiences (success/failure) and rehabilitation duration (1 to 5 years). An a priori decision was made to limit follow-up from 1 to 5 years, because outcomes of pain and physical functioning from 1 to 5 years following TKA are reasonably stable.[13-15] Overall, patients were included with diversity in age and sex. Patients were approached via three different hospitals. Physicians working in these three hospitals were asked to invite eligible patients. Eligibility criteria for patient selection can be found in Table 1. Interviews with Dutch patients were held at their homes. Patients in Belgium were interviewed during individual online meetings as a consequence of the regulations during the COVID-19 pandemic. Interviews were audio-recorded, and additional field notes were made during and after the interviews.

Table 1. Eligibility criteria by patient subgroup

Subgroup 1: Patients with knee OA without surgical indication
- Patients with self-reported knee OA or knee pain (for more than 3 months)
- 18 years or older
- Dutch as native language
- No surgery for contralateral TKA
- No hearing or speech impairment
- Able and willing to participate and provide informed consent

Subgroup 2: Patients with knee OA scheduled for TKA
- Patient with a clinical diagnosis of knee OA and scheduled for TKA in one of the three participating hospitals
- 18 years or older
- Dutch as native language
- No surgery for contralateral TKA
- No hearing or speech impairment
- Able and willing to participate and provide informed consent
- Dutch patients: living within 50 kilometres from one of the two participating hospitals
- Belgian patients: willing to be present at UZ Gent at the day of the interview

Subgroup 3: Patients 1 to 5 year after TKA
- Patients with a primary TKA for 1 to 5 years
- Surgery performed in one of the three participating hospitals
- 18 years or older
- Dutch as native language
- No hearing or speech impairment
- Able and willing to participate and provide informed consent
- Dutch patients: living within 50 kilometres from one of the two participating hospitals
- Belgian patients: willing to be present at UZ Gent at the day of the interview

Interviews with knee specialists

Through purposive sampling we included knee specialists as 1) orthopaedic surgeons, performing at least 30 primary TKA procedures a year, 2) physiotherapists (specialized in knee rehabilitation) involved in the care of at least 10 TKA patients a year and 3) orthopaedic PAs and NPs (both in the Netherlands) with at least 50% of their patient contacts with TKA patients. Orthopaedic knee specialists with Dutch as native language, working in the Netherlands or in Belgium were recruited through the research team. Participants were invited via email or telephone to participate and were also asked to nominate potential other knee specialists (snowball-sampling).[16] Within the sample of knee specialists we sought variety in current working environment (university-, general- or specialized hospital and physiotherapy practice), working experience in TKA surgery and care. Knee specialists were interviewed face-to-face at their hospital office, physiotherapy practice or during a conference.

Data collection

An interview guide was developed based on a review of the literature on poor outcome after TKA and clinical expertise of the research team (Supplementary Table S2).[6] The questions had an open-ended format and were adapted to the specific subgroup of participants (Table 2). The interview guides were discussed with the patient research partners and then pilot-tested in one postoperative patient and one orthopaedic surgeon, leading to minor changes in the wording of interview guides. All interviews were conducted between May and November 2020 by one researcher (MtM, PhD student), who had formal interview training and had no pre-existing relationship with any of the participants. The interviewer had previous experience in working with TKA patients as research nurse and as researcher. Before the interviews started, participants were asked to fill out a short questionnaire to collect participants' characteristics. A summary of the interview was sent to the interviewee after each interview as a member check to assure data validity. Two patients responded and confirmed that they recognised their experiences in the summaries.

Data collection ended after twenty-five patients interviews and fifteen interviews with knee specialists as data saturation had been reached (no new information emerged from the last 2 interviews).

Table 2. Main topics of patients' interview guide with one example question and some probing questions

Topic	Example question
Outcome of TKA	What are important outcomes of TKA for you? • Why are these important outcomes for you? • Could you tell me more about this?
Expectations of the outcome	To what extent are your expectations regarding TKA outcome been fulfilled? • Why are these expectations not fulfilled? • How realistic were these expectations?
Less successful outcome	When do you consider the result of TKA less successful? • Why is the result less successful for you?
Unsuccessful outcome	When do you consider the result of TKA unsuccessful? • Why is the result unsuccessful for you? • What do you think is the worst-case scenario?
Dissatisfaction with the outcome	What would be reasons for you to be dissatisfied with your TKA? • Which factors play a role in this? • Could you explain to me?
Timepoint for outcome	What is for you the optimal time after surgery to assess the success of TKA? • How do you determine if the prosthetic knee is not working for you? • Could you explain to me?

Data analysis

Audio recordings were transcribed by a commercial third party (Secretaresse Hulp), anonymized, and checked for accuracy against the original audio recordings. Transcripts were analysed using Atlas.ti vV.8. Data for patients and knee specialists were analysed separately. Following the grounded theory approach with thematic analysis, coding was performed in 3 steps: open, axial, and selective coding.[10] The first step started with reading and re-reading the transcripts for familiarization. Relevant fragments were selected in the interviews and each fragment was given a label (open coding). Second, these open codes were categorized (axial coding). From these axial codes the core themes were identified (selective coding). To support the coding process, field notes were made during the interviews. Data collection and data analysis was continuously alternating in a cyclic process. To enhance trustworthiness[17], the first three interviews in each group (patients and knee specialists) were independently coded by two researchers (JV and MtM). The remaining interviews were coded by one researcher (MtM). Throughout this process, three researchers (JV, CvdE and MtM) continuously and repetitively reflected on, compared, discussed, refined and adjusted the codes in order to carefully determine the number and wording of themes in an iterative design. The identified themes were thoroughly discussed until consensus was reached in the research team (comprising a Dutch orthopaedic surgeon, a Belgian orthopaedic surgeon, a psychologist, a physiotherapist, and a nurse). Finally, quotes were extracted that related to the subthemes. Quotes were translated into English in collaboration with a professional translator.

Ethics

This study was conducted in accordance with the ethical standards in the 1964 Declaration of Helsinki. The ethical Review Board of the Radboud University Medical Centre, Nijmegen in the Netherlands exempted the study (ref. number: 2019/5283) from ethical approval according to the Medical Research Involving Human Subjects Act. In Belgium, the institutional ethics committee of the Gent University Hospital granted approval (BC-07096). All participants gave informed consent.

Results

Participant characteristics

In Tables 3 and 4, respectively, the characteristics of the patient and knee specialist samples are shown. The duration of the interviews varied from 25 to 85 minutes (mean (SD): 51.2 (11.8) minutes) in patients and from 25 to 59 minutes (mean (SD): 44.7 (8.7) minutes) in knee specialists.

Table 3. Characteristics and clinical details of Dutch and Belgian patients

Characteristics	The Netherlands	Belgium
Interviews, n	18	7
Subgroup, n		
Patients with knee OA without surgical indication	2	NA
Patients with knee OA scheduled for TKA	8	3
Patients 1 to 5 year after TKA	8	4
Age in years, median (25^{th} – 75^{th} percentile)	64 (62-69)	61 (57-66)
Woman, n	12	4
Level of education[a], n		
Primary & Secondary	10	6
Tertiary	8	1
Currently employed, n	4	0
General health score[b], median (25^{th} – 75^{th} percentile)	47 (29-53)	43 (34-52)
Pain [c], median (25^{th} – 75^{th} percentile)	38 (26-53)	36 (27-46)
Physical functioning [d], median (25^{th} – 75^{th} percentile)	55 (39-60)	54 (38-62)

[a] Primary: primary education; Secondary: lower secondary education, upper secondary education; Tertiary: short-cycle tertiary education, bachelor's or equivalent, master's or equivalent.
[b] Position marked on visual analogue scale (100 mm) from best health (left side: 0/100) to worst health (right side: 100/100).
[c] Position marked on visual analogue scale (100 mm) from no pain (left side: 0/100) to worst pain (right side: 100/100).
[d] Position marked on visual analogue scale (100 mm) from no problem (left side: 0/100) to much difficulty (right side: 100/100).

Main results

Four themes and 15 subthemes were identified (Table 5). Each theme is described in detail below, separately for both groups of participants. In Table 6, quotations from patients and knee specialists are displayed for each subtheme.

Table 4. Characteristics and working experience of Dutch and Belgian knee specialists

Characteristics	The Netherlands	Belgium
Interviews, n	12	3
Orthopedic surgeons, n	**6**	**3**
General hospital	3	2
Specialized hospital	2	NA
University hospital	1	1
Physician assistants/Nurse practitioner, n	**2**	NA
General hospital	2	
Physiotherapists, n	**4**	NA
General hospital	1	
Physiotherapy practice	3	
Experience in TKA surgery, treatment and/or rehabilitation in years, median ($25^{th}-75^{th}$ percentile)	10 (7-12)	16 (14-22)
Number of performed primary TKA surgeries per year, median ($25^{th}-75^{th}$ percentile)	100 (24-185)	100 (65-100)

Table 5. Themes and subthemes identified in thematic analysis

Theme	Subtheme	Apply to
Lingering pain	Pain during the first postoperative period	patients
	Pain medication to ease the pain	patients
	Pain flares during and/or after physical activity	patients
	Continued unacceptable pain to the patient	knee specialists
Stagnating mobility	Limitations in physical activities	patients
	Not fulfilling conditions for mobility	knee specialists
	Limited walking ability	knee specialists
Complications and revision surgery	Revision surgery	patients & knee specialists
	Complications	patients & knee specialists
	Surgical technical limitations	knee specialists
Getting used to the prosthetic knee	Disappointing first postoperative period	patients
	Movement anxiety	patients
	Lingering pain creates uncertainty	patients
	Unhelpful thoughts	patients
	Awareness of the prosthesis	patients

Lingering pain

Negative outcome expectations for preoperative patients were no improvement in pain throughout the day and night and pain that limits patients in resuming valued activities. Postoperative patients reported that the typical preoperative OA pain had disappeared postoperatively, but some patients reported that it had been replaced by a new, different type of pain. Most patients experienced this new type of pain during the first postoperative period (6 to 8 weeks), some of them had a lot of pain or a burning sensation in the knee and could not sleep at night while others described the pain as muscle ache or nagging pain. Some patients did not get adequate pain relief because they discontinued taking pain medication because of side effects (e.g. nausea or vomiting) while other patients preferred not to use any pain medication other than paracetamol. Up to a year a number of patients experienced pain flares during and/or after physical activity such as climbing stairs, walking long distances (e.g. one hour), and during and/or after more strenuous exercise such as hiking in the mountains or sports (e.g. golf, fitness). Patients described these pain flares as an irritating or cramping sensation and interpreted this as a signal for taking rest for the remainder of the day.

Knee specialists reported that they have concerns when patients continue to have a lot of pain, beyond an acceptable level for the patients or when the cause of pain remained elusive. These patients are seen more often at the outpatient clinic and sometimes receive additional treatment (i.e. additional follow-up consults or guidance from the pain specialist). Knee specialists sometimes felt that nothing could be done for the patients.

Stagnating mobility

Most negative outcome expectations and experiences of both pre- and postoperative patients were focused on limitations during physical functioning and in resuming valued activities. In addition, preoperative patients mentioned a decline in walking ability and continued reliance on a cane, crutch or walker as negative outcome expectations. Postoperative patients described poor function of the knee in terms of a tight feeling in the knee, a feeling of knee stiffness and an unreliable knee. They felt hampered in their mobility, and consequently in their activities of daily living and leisure activities such as climbing stairs, walking long distances, gardening, cycling, playing tennis, or shopping. Patients indicated the inability to resume valued activities in and around the house as an adverse consequence of TKA. Some patients found it disappointing that they could not return to an active lifestyle.

Negative outcomes for knee specialists were non-fulfilling conditions that hamper patients in their mobility. Knee specialists considered an extension or flexion limitation, (mid-flexion) instability, or stiffness of the knee as a negative outcome. Particularly, when an extension limitation affects the patient's ability to walk and made walking tiresome. In patients with stiffness within 3 months post-TKA, the patient's knee needs to be manipulated under anesthesia (MUA) with the purpose of regain full range of motion. In addition, knee specialists and especially physiotherapists were concerned when patients remain limited in their walking ability; that is when they are not able to walk independently for a short distance or when they are not able to walk without pain or discomfort.

Complications and revision surgery

Some patients experienced surgical complications such as a swollen knee, a chronic inflammation of the knee, thrombosis, stiffness of the knee or a prosthesis infection that resulted in a negative experience after TKA. Some of the study patients experienced a manipulation under anaesthesia (MUA) or revision surgery, i.e. debridement, antibiotics and implant retention (DAIR) because of infection.

Knee specialists mentioned several complications that might contribute to a negative experience for patients: swollen knee, reactive knee (redness, swelling, heat), deep venous thrombosis, wound healing disorders, vascular injury, severe stiffness, aseptic loosening, infection and malrotation of the prosthesis. Early revision surgery is a clear indication of poor response to TKA according to knee specialists.

In addition, orthopaedic surgeons mentioned surgical technical shortcomings, such as malposition, prosthetic loosening, malalignment, and mechanical failures, contributing to a negative experience for the orthopaedic surgeon.

Getting used to the prosthetic knee

For most patients, the first postoperative period was hard, exhausting, disappointing, and sometimes with emotional impact due to unexpected pain, surgical complications, medication side effects, and being dependent on help from other people. Patients worried whether their pain was normal and comparable to other patients and whether the knee was recovering properly. Some patients who experienced a difficult recovery after the first postoperative period and the ones with persisting pain after 6 to 9 months experienced uncertainty on several domains, some experienced movement anxiety. They worried about the future, slept poorly, and were distressed. Patients had struggles with adjusting to their prosthetic knee in daily life. One patient took early retirement while others were able to return to work but later than initially planned and/or with temporary adjustments. Adaptations (in duration, frequency, bracing and other support measurements) related to movement and sport were mentioned or more in general accepting that not everything is possible with the prosthetic knee and adjusting their level of activity. Some patients struggled with unhelpful thoughts that limited their motivation to practice physiotherapy exercises and their hope that the knee would become better. Other patients were eager to engage in different tasks despite their pain and limited knee function and refused to be negatively affected by their prosthetic knee. More in general, several patients mentioned that being aware of the prosthesis all the time was an unexpected, unpleasant experience. Preoperative patients expected to be recovered within 3 to 6 months, while most postoperative patients indicated that they were recovered in 6 to 7 months, but for some it took longer. Two patients indicated that they had fully recovered only after 2 years, while knee specialists indicated that patients should take into account a rehabilitation duration of one year.

Table 6. Subthemes and quotations

Subtheme	[a]	Quotation
Pain during the first postoperative period	Preoperative patient #5, 67	I would not be satisfied if the pain has not decreased or remained the same.
	Postoperative patient #12, 62	I slept poorly during the first few weeks, purely because of the new knee. It felt like a burning sensation. Every time I wanted to turn over in bed, I woke up, and then it took me a while to find a comfortable position again. Those nights were actually the most disappointing for me.
Pain medication to ease the pain	Postoperative patient #21, 68	I couldn't tolerate the pain medication after the surgery. As a result, I relied on paracetamol during the first few weeks. It was terrible because the pain was unbearable.
Pain flares during and/or after physical activity	Preoperative patient #3, 59	If pain prevents me from resuming my daily activities, I would not be happy. I just want to be able to go up and down the stairs with a laundry basket and take care of the garden myself.
	Postoperative patient #10, 46	After about an hour at the gym, or after an hour of swimming, I start experiencing a sense of irritation, indicating that I need to stop. At that point, I prefer to sit down with my legs up, and take a moment of rest for my knee.
Continued unacceptable pain to the patient	Knee specialist #10, nurse practitioner	These patients have a lot of pain all the time and if you want to bend their knee, everything hurts. I find that worrisome! Then I think TKA might have been a bad choice.
Limitations in physical activities	Preoperative patient #14, 48	In the past few years, I already had to limit my activities. If I experience even greater limitations in my activities after the operation, then the knee prosthesis has been of no use to me. I would like to resume my work in a department store, and it would be nice if I could go out with my children again.
	Postoperative patient #7, 67	I constantly feel like I'm going through my knee, that makes it difficult to rely on my knee. Therefore, I don't dare to climb stairs or to go out for a long walk.
Not fulfilling conditions for mobility	Knee specialist #14, orthopaedic surgeon	After 8 weeks, I want patients to reach 110 degrees of flexion and 0 degrees of extension. If patients have less than 90 degrees of flexion and/or 10 degrees or more of extension at that point, I am not satisfied.
Limited walking ability	Knee specialist #2, physiotherapist	Training the quadriceps function and gait pattern is crucial. If patients start walking longer distances with an abnormal gait pattern, they will experience difficulties. Moreover, learning a new gait pattern can be challenging.
Complications	Postoperative patient #17, 62	The knee went red, warm and swollen. I started with antibiotics because there was an infection going on.
	Knee specialist #15, orthopaedic surgeon	There are mild and severe complications, an aseptic loosening is an example of a serious complication.

Revision surgery	Knee OA patient without surgical indication, #2, 72	The surgeon said that the prosthesis could have a lifespan of 15 to 20 years. So if a reoperation is required within 10 years, I would be disappointed.
	Postoperative patient #11, 64	I was able to flex the knee up to 70 degrees, and there was no further progress. I waited for another 4 weeks, but I didn't make any progress, not even a millimetre. The extension of the knee also deteriorated over time, leading to the decision to manipulate the knee under anaesthesia.
	Knee specialist #7, physiotherapist	Early revision surgery as a consequence of for example arthrofibrosis or prosthetic loosening, is perhaps one of the most objective measures of failure.
Disappointed first postoperative period	Postoperative patient #6, 72	The first 3 months really disappointed me. I thought: "a new knee, wound healed and done!" However, that wasn't the case. We cooled the knee with ice for at least 6 to 7 weeks.
Movement anxiety	Postoperative patient #8, 64	Exercising was scary because I was anxious to bend the knee. I was afraid that something would tear. I had to trust my knee and allow it to relax in order to achieve further flexion. It was ultimately a matter of building confidence.
Lingering pain creates uncertainty	Postoperative patient #10, 46	When the pain persists for such a long time and the physiotherapist refers you back to the surgeon, it can create a sense of insecurity. You start to wonder if the pain will ever go away.
Unhelpful thoughts	Postoperative patient #11, 64	There is progress happening, and that's what keeps me going. It's my mindset, and it's important. You almost need psychological support to avoid falling into a slump when things are going so poorly.
Awareness of the prosthesis	Postoperative patient #16, 57	I didn't expect to feel the prosthesis every day. This possible experience was not communicated to me prior to the operation. While it is possible to adapt and live with it, this sensation was unexpected for me.

[a] Patient participant indexed by the #patient identification number and age or
Knee specialist participant indexed by the #knee specialists identification number and profession;

Discussion

Findings of this study highlight that knee specialists put emphasis on surgical failure, unexplained pain, limited walking ability and impairments that limit patients' physical functioning, while patients were focused on the arduous process of getting used to the prosthesis, lingering pain, awareness of the artificial knee and limitations they experience during valued and daily activities.

In line with previous research[18,19], our study showed that the process of getting used to the prosthesis, the experiences of adjusting physically and mentally to the prosthesis, were in top of mind of patients. Especially in patients for whom rehabilitation took longer than anticipated, this process of adapting to their artificial knee was dominant in their stories. Patients for whom additional efforts for improvement did not result in their expected outcome, expressed deep frustration with what they perceived to be a lack of adequate guidance or help from the health care providers. On the other hand, knee specialists felt that the process of getting used to the prosthesis in their stories was part of the postoperative process and did not explicitly acknowledge that the process of recovery can contribute to a negative experience. These differences in views between patients and knee specialists might contribute to discrepancies in their perception of poor response to TKA.

Patients' experiences of lingering pain and limitations in performing valued and daily activities were also contributing to the negative appraisal of the TKA procedure. Comparable results

were found in previous studies that focused on asking patients what results matter the most to patients undergoing a knee or hip replacement.[9,20,21] Patients ranked three outcomes as their highest priorities: pain relief, functional recovery, and improved quality of life.[20] Another study by Whitebird et al. identified the ability to walk without pain or discomfort, pain relief, and returning to an active lifestyle as important outcomes.[21] While both studies focused on pain relief, many patients specifically discussed pain in relation to specific activities like mobility and walking. The association between pain and performing valued activities is also reflected in our findings as patients reported experiences of pain flares during and/or after physical activity. Knee specialists in our study tend to put more emphasis on surgical failure, unexplained pain, limited walking ability and impairments that limit patients' physical functioning. Remarkably, surgical failure is not incorporated in any of the definitions for poor response after TKA used in the literature.[6] Our findings indicate that, besides complications and revision surgery, lingering pain, limitations in walking ability and the ability to perform (valued) activities of daily living and/or work are relevant outcome domains for measuring poor response to TKA according to both patients and knee specialists.

There is evidence that TKA patients tend to have overly high expectations going into surgery.[22,23] Knee specialists in our study confirmed that some patients persist in unrealistic expectations on the outcome of a TKA, even after comprehensive preoperative consultation. Most knee specialists in our study mentioned that they discuss these (unrealistic) expectations with their patients. Nevertheless, it is important to encourage patients to list what they would like or expect to do post-TKA.[24] Knee specialists must appropriately council them regarding the relative probability that they would be able to accomplish each of their stated goals.[24] In case of a discrepancy between what the patient expects and what knee specialists know TKA can deliver, the first step is for the knee specialist to explain how realistic the patients' expectations are.[24] This should be seen as an essential component of preoperative consultation.

These findings have several important clinical and research implications. Our findings can inform shared decision making for TKA. We found that lingering pain, impaired mobility and the inability to resume valued activities are important adverse consequences of TKA. Preoperative consultation must therefore include a discussion on the likelihood of those outcomes. Furthermore, our study provided a full picture on the variety of potential adverse consequences of TKA that could contribute to a negative appraisal of its effect. However, we did not identify the relative importance of these consequences. Future research should focus on the prioritization of adverse consequences of TKA for patients with OA that may contribute to poor response, from both the perspective of patients and knee specialists.

One of the strengths of this study is that, to our knowledge, this is the first study investigating expectations and experience of patients and perceptions of knee specialists on adverse consequences that may contribute to the negative appraisal of the TKA procedure. Another strength is that the interviews were conducted by a PhD candidate who did not have a relationship with the interviewed patients prior to the study, reducing the risk of response bias. Furthermore, purposive sampling and participant recruitment from two hospitals in the Netherlands and one in Belgium allowed the inclusion of a wide variety of participants, leading through a thorough evaluation of all possible experiences that may contribute to the negative appraisal of the TKA procedure. Another strength is the involvement of patient research partners in the study design.

Potential limitations must be considered while interpreting the findings. Firstly, our results describe a process that unfolds over time, data were collected at one time point and thus, for postoperative patients, relied on the participants' recall of their TKA journey. However, some patients indicated they could now better reflect on that period than they could during the rehabilitation period. We attempted to minimize recall and salience bias by asking patients about their own experiences, about previously mentioned experiences of other participants and by asking probing questions about all sort of details. Secondly, signs of nonverbal communication during the interviews with patients in Belgium could have been missed because these interviews were online as a consequence of the regulations during the COVID-19 pandemic. However, we utilized video conferencing software which is seen as the closest to the gold standard of interviewing.[25]

Thirdly, only patients and knee specialists who were able to communicate (read and speak) in Dutch were included to ensure that all interviews could be conducted in the native language. Thus, cultural differences, and different health care systems can make these results less generalizable.

In conclusion, our study provides a comprehensive overview of potential adverse consequences from the perspective of both patients and knee specialists. Our findings highlight that knee specialists put more emphasis on surgical failure, unexplained pain, limited walking ability and impairments that limit patients' physical functioning, while patients' experiences were more focused on the arduous process of getting used to the prosthesis, lingering pain, awareness of the artificial knee and limitations they experience during valued and daily activities. Aspects associated with the difficult process of adapting to the prosthesis need to be addressed during shared decision making.

Acknowledgements

The authors are grateful to the patients and knee specialists who were willing to participate in this research, and thank nurse practitioner Keetie Kremers for her efforts in patient recruitment in one of the three participating hospitals. The authors also thank Manorma Dwarkasing and John van de Langenberg (patient research partners) for their contribution to the study.

References

1. Kamaruzaman H, Kinghorn P, Oppong R. Cost-effectiveness of surgical interventions for the management of osteoarthritis: a systematic review of the literature. BMC Musculoskelet Disord 2017;18(1):183.
2. Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. BMJ open 2012;2(1):1-12.
3. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010;468(1):57-63.
4. Gunaratne R, Pratt DN, Banda J, Fick DP, Khan RJK, Robertson BW. Patient Dissatisfaction Following Total Knee Arthroplasty: A Systematic Review of the Literature. J Arthroplasty 2017;32(12):3854-3860.
5. Halawi MJ, Jongbloed W, Baron S, Savoy L, Williams VJ, Cote MP. Patient Dissatisfaction After Primary Total Joint Arthroplasty: The Patient Perspective. J Arthroplasty 2019;34(6):1093-1096.
6. Te Molder MEM, Smolders JMH, Heesterbeek PJC, van den Ende CHM. Definitions of poor outcome after total knee arthroplasty: an inventory review. BMC Musculoskelet Disord 2020;21(1):378.
7. Noble PC, Fuller-Lafreniere S, Meftah M, Dwyer MK. Challenges in outcome measurement: discrepancies between patient and provider definitions of success. Clin Orthop Relat Res 2013;471(11):3437-45.
8. Jacobs CA, Christensen CP. Factors influencing patient satisfaction two to five years after primary total knee arthroplasty. J Arthroplasty 2014;29(6):1189-91.
9. Singh JA, Mehta B, Mirza SZ, et al. When Has a Knee or Hip Replacement Failed? A Patient Perspective. J Rheumatol 2021;48(3):447-453.
10. Charmaz K. Constructing grounded theory. 2nd ed. Saga; 2014.
11. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3:77-101.
12. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014;89(9):1245-51.
13. Fortin PR, Penrod JR, Clarke AE, et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. Arthritis Rheum 2002;46(12):3327-30.
14. Lingard EA, Katz JN, Wright EA, Sledge CB. Predicting the outcome of total knee arthroplasty. J Bone Joint Surg Am 2004;86(10):2179-86.
15. Nilsdotter AK, Toksvig-Larsen S, Roos EM. A 5 year prospective study of patient-relevant outcomes after total knee replacement. Osteoarthritis Cartilage 2009;17(5):601-6.
16. Green J, Thorogood N. Qualitative methods for health research. 3rd ed. ed. SAGE; 2014.
17. Nowell LS, Norris JM, White DE, Moules NJ. Thematic Analysis: Striving to meet the thrustworthiness criteria. International Journal of Qualitative Methods 2017;16(1):1-13.
18. Randall E, Bryan S, Black C, Goldsmith LJ. What matters to patients following total knee arthroplasty? A grounded theory of adapting to a knee replacement. BMC Musculoskelet Disord 2022;23(1):845.
19. Woolhead GM, Donovan JL, Dieppe PA. Outcomes of total knee replacement: a qualitative study. Rheumatology (Oxford) 2005;44(8):1032-7.
20. Goodman SM, Mehta B, Mirza SZ, et al. Patients' perspectives of outcomes after total knee and total hip arthroplasty: a nominal group study. BMC Rheumatol 2020;4:3.
21. Whitebird RR, Solberg LI, Ziegenfuss JY, et al. Personalized outcomes for hip and knee replacement: the patients point of view. J Patient Rep Outcomes 2021;5(1):116.
22. Mannion AF, Kämpfen S, Munzinger U, Kramers-de Quervain I. The role of patient expectations in predicting outcome after total knee arthroplasty. Arthritis research & therapy 2009;11(5):R139.
23. Nilsdotter AK, Toksvig-Larsen S, Roos EM. Knee arthroplasty: are patients' expectations fulfilled? A prospective study of pain and function in 102 patients with 5-year follow-up. Acta orthopaedica 2009;80(1):55-61.

24. Dunbar MJ, Richardson G, Robertsson O. I can't get no satisfaction after my total knee replacement: rhymes and reasons. Bone Joint J 2013;95-b(11 Suppl A):148-52.
25. Thunberg S, Arnell L. Pioneering the use of technologies in qualitative research - A research review of the use of digital interviews. International Journal of Social Research Methodology 2022;25(6):757-768.

Supplementary material

Supplementary Table S1: SRQR checklist

Standards for Reporting Qualitative Research (SRQR)*

http://www.equator-network.org/reporting-guidelines/srqr/

	Page/line no(s).
Title and abstract	
Title - Concise description of the nature and topic of the study Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended	1
Abstract - Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	2
Introduction	
Problem formulation - Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	3
Purpose or research question - Purpose of the study and specific objectives or questions	3
Methods	
Qualitative approach and research paradigm - Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/ interpretivist) is also recommended; rationale**	5
Researcher characteristics and reflexivity - Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	6
Context - Setting/site and salient contextual factors; rationale**	5-6
Sampling strategy - How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale**	5-6
Ethical issues pertaining to human subjects - Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	7-8
Data collection methods - Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale**	6-7
Data collection instruments and technologies - Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	6-7

Units of study - Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	Tables 3 and 4
Data processing - Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/de-identification of excerpts	7
Data analysis - Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale**	7
Techniques to enhance trustworthiness - Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale**	7

Results/findings

Synthesis and interpretation - Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	8-12
Links to empirical data - Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	Tables 5 and 6

Discussion

Integration with prior work, implications, transferability, and contribution(s) to the field - Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/generalizability; identification of unique contribution(s) to scholarship in a discipline or field	13-16
Limitations - Trustworthiness and limitations of findings	15

Other

Conflicts of interest - Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	17
Funding - Sources of funding and other support; role of funders in data collection, interpretation, and reporting	17

*The authors created the SRQR by searching the literature to identify guidelines, reporting standards, and critical appraisal criteria for qualitative research; reviewing the reference lists of retrieved sources; and contacting experts to gain feedback. The SRQR aims to improve the transparency of all aspects of qualitative research by providing clear standards for reporting qualitative research.

**The rationale should briefly discuss the justification for choosing that theory, approach, method, or technique rather than other options available, the assumptions and limitations implicit in those choices, and how those choices influence study conclusions and transferability. As appropriate, the rationale for several items might be discussed together.

Supplementary Table S2: Interview guide knee specialists

Introduction

Total knee arthroplasty (TKA) is considered to be a cost-effective intervention for the treatment of advanced knee osteoarthritis (OA). However, increasing evidence shows that a proportion of patients (around 20%) could be considered as a poor responder to TKA (for example show no or too little improvement) in terms of chronic knee pain, functional disability, poor quality of life, and dissatisfaction after TKA. Moreover a successful outcome according to the physician is not a guarantee for treatment success as perceived by patients. This may be caused by different perceptions of adverse consequences after TKA. With this interview study we are hoping to learn about adverse consequences after TKA, and differences between patients and knee specialists.

1. Could you briefly tell me about your position as knee specialist?
   Probes:
   1.2 How much time do you spend on primary TKA procedures?
   1.3 Do you organize information meetings about TKA procedures?

2. What are important outcomes of TKA for you as knee specialist?
   Probes:
   2.1 Why are these important outcomes for you as knee specialist?
   2.2 What are important outcomes for patients do you think?

3. When are your expectations regarding TKA outcome been fulfilled?
   Probes:
   3.1 When and why are expectations not fulfilled?

4. When do you consider the result of the TKA less successful for the patient?
   Probes:
   4.1 Why is the result less successful for the patient?
   4.2 When is the result less successful for you as a knee specialist?
   4.3 Why is the result less successful for you?

5. When do you consider the result of the TKA unsuccessful for the patient?
   Probes:
   5.1 Why is the result unsuccessful for the patient?
   5.2 When is the result unsuccessful for you as a knee specialist?
   5.3 Why is the result unsuccessful for you?

6. What are reasons for patients to be dissatisfied with the TKA?
   Probes:
   6.1 Which factors play a role in this for patients?
   6.2 Could you explain to me?

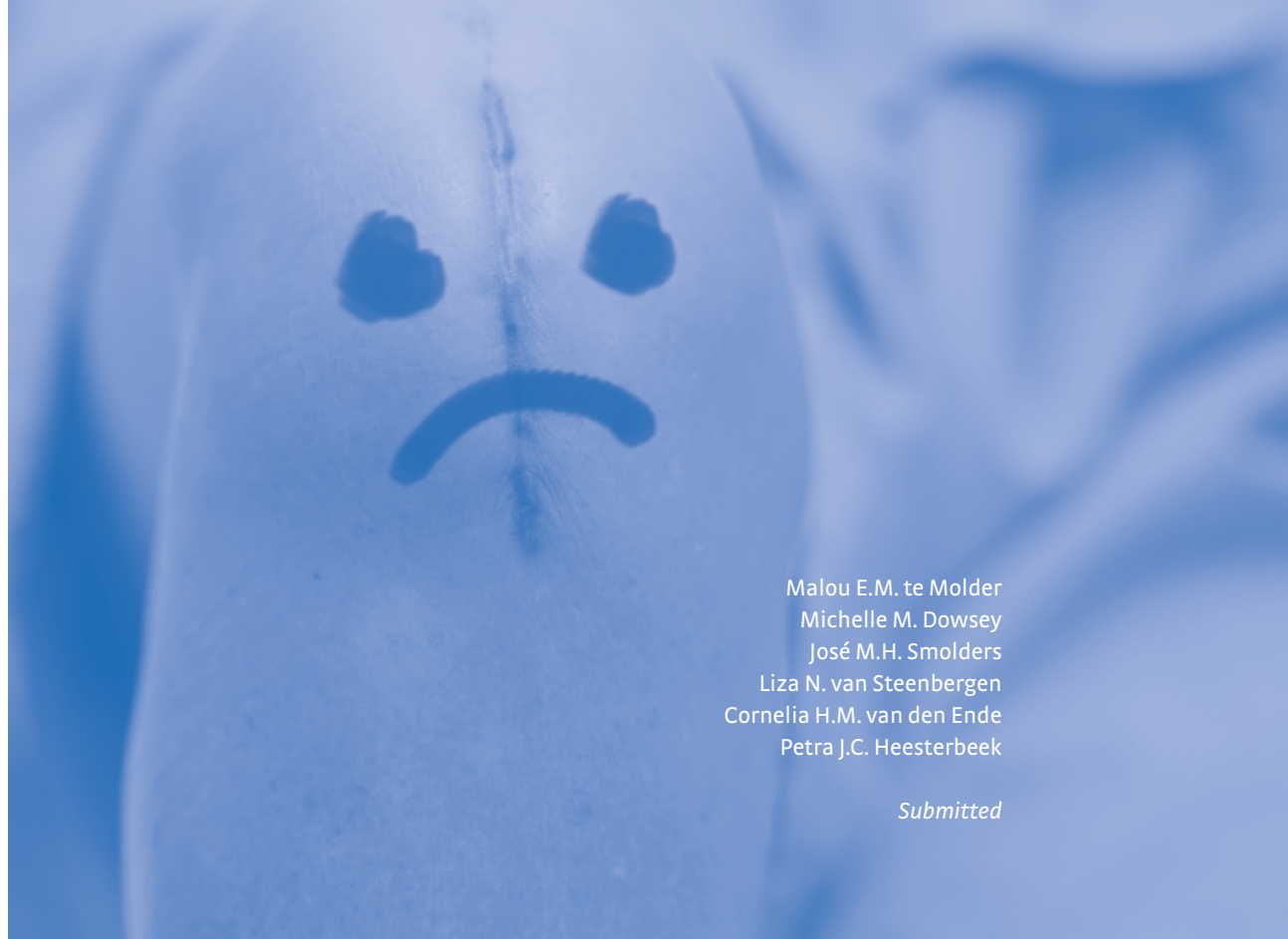
7. What are reasons for you as a knee specialist to be dissatisfied with the TKA?
   Probes:
   7.1 Which factors play a role in this for you?

8. What is for you as a knee specialist the optimal time after surgery to assess the success of TKA?
   Probes:
   8.1 How do you determine if the prosthetic knee is successful or not?
   8.2 Could you explain to me?

Thank you for your time and sharing your perceptions. We will send you a summary of our interview and we would ask you to check the summary. If something has not been interpreted correctly or if you do not agree with something in the summary, we ask you to respond. We hope this information will help to identify what adverse consequences of TKA contribute to a poor response to TKA according to the perspective of patients and knee specialists.

Chapter 5

Comparison of performance of different definitions of poor response after total knee arthroplasty using the Dutch Arthroplasty Register and the Osteoarthritis Initiative database



Malou E.M. te Molder
Michelle M. Dowsey
José M.H. Smolders
Liza N. van Steenbergen
Cornelia H.M. van den Ende
Petra J.C. Heesterbeek

Submitted

Abstract

Purpose

A variety of definitions of poor response to total knee arthroplasty (TKA) is used in the literature, which impedes the comparisons of response after TKA over time and across hospitals and countries. The aim of this study was to compare the prevalence, overlap and discriminative accuracy of 15 definitions of poor response after TKA using two large databases.

Methods

Data of patients one year after primary TKA from the Dutch Arthroplasty Register (LROI) (n=12,275) and the Osteoarthritis Initiative (OAI) database (n=204) were used to examine the prevalence, overlap (estimated by Cohen's kappa) and discriminative accuracy (sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and Youden index) of 15 different definitions of poor response after TKA. In the absence of a true gold standard for measuring poor responder to TKA, the numeric rating scale (NRS) satisfaction (0-10; dichotomized: >6 'responder' vs ≤6 'poor responder') and the Global assessment of Knee impact (0-10; dichotomized: <4 'responder' vs ≥4 'poor responder') were used as anchor for the assessment of discriminative accuracy for the LROI dataset and OAI dataset, respectively. 12 definitions could be tested in the LROI dataset and 7 in the OAI dataset.

Results

The median (25^{th} – 75^{th} percentile) prevalence of poor responders of the examined definitions was 18.5% (14.0%-25.5%) and median Cohen's kappa for the overlap between pairs of definitions was 0.41 (0.32-0.59). Median (25^{th} – 75^{th} percentile) sensitivity was 0.45 (0.39-0.54), specificity was 0.86 (0.82-0.94), PPV was 0.45 (0.34 – 0.62), NPV was 0.89 (0.87 – 0.89) and the Youden index was 0.36 (0.20-0.43).

Conclusions

The prevalence of poor responders was discordant across the different definitions in this study, and the majority of pairs of definitions showed a fair or moderate agreement suggesting a lack of overlap between different definitions of poor response to TKA. None of the definitions we examined adequately classified poor responders of TKA. In contrast, absence of poor response could be classified with confidence. A future unambiguous definition of poor response to TKA is needed and should preferably include multiple outcome domains and combine a relative and absolute change.

Keywords

Total knee arthroplasty, Patient-reported outcome measurement, Poor response, Discriminative accuracy

Introduction

Total knee arthroplasty (TKA) is an effective surgical procedure for patients with advanced symptomatic knee osteoarthritis (OA). Nonetheless, a substantial proportion of patients experience a poor response to TKA (i.e. show no or little improvement) in terms of chronic knee pain[1], functional disability[2], poor quality of life[3], and/or dissatisfaction after TKA.[4,5]

In the literature, it is generally assumed that the proportion of patients with poor response to TKA is about 15 to 20%.[1,4,5] However, a variety of dichotomous definitions of poor response to TKA comprising one or more different dimensions of outcome are used in literature[6] to quantify the proportion of patients with poor response to TKA. These definitions were used in small-scale, single institution studies at different time points (1 month and 5 years), and based on different outcome dimensions (e.g. pain, daily functioning, knee function, satisfaction).[6] The variety of definitions used impedes the comparisons of poor response to TKA over time and across hospitals and countries, because definitions cannot be exchanged and results are difficult to compare. These findings also illustrate the lack of international consensus on the definition of "poor response" to TKA, although the need to use a combined endpoint has been recognized to accurately describe failure (i.e. poor response) after TKA.[7]

Insight in the prevalence of poor responders, overlap of definitions and discriminative accuracy in databases with large patient samples and large sets of patient-reported outcome measures (PROM) scores, provides the opportunity to compare the performance of different definitions of poor response to TKA. Information on the performance of definitions can serve as input for further research to reach consensus on the definition for poor response. Furthermore, it provides insight in whether definitions measure different outcome domains or whether definitions might be interchangeable. Therefore, the purpose of this study was to examine the performance of a set of definitions, derived from literature and newly composed by an expert group, defining poor response to TKA in existing databases.

Materials and methods

Data Sources

This study used data from the Dutch Arthroplasty Register (LROI) and the Osteoarthritis Initiative (OAI) database.

The LROI is a nationwide, population-based register of joint arthroplasties that includes information on >950,000 joint arthroplasties in the Netherlands since 2007.[8] It has a coverage of 100% of Dutch hospitals and a completeness up to 99% for primary TKAs.[8] Patient-reported outcome measures (PROMs) for TKAs were registered in the LROI since 2015. The response rate for matched preoperative and 12-month postoperative PROMs in the period 2015-2020 was 38%.[9]

The OAI database is a multicentre, longitudinal cohort study which includes 4796 patients with, or at risk of developing, symptomatic OA of the knee. Patients were enrolled between February 2004 and May 2006 in four centres and were followed for nine years. The OAI database is available for public access at (https://data-archive.nimh.gov/oai/), and details have been published elsewhere.[10]

Ethical approval

All data in the LROI register were registered as part of routine clinical care, and the present study placed no additional burden on the patient. Data used from the public OAI database are not proprietary. Ethical approval for collecting information about the subjects was provided by the OAI, and informed consent was obtained from all individual participants included in the study. Investigators had access only to unidentifiable patient data in both the LROI register and the OAI database. Therefore, no ethical approval was necessary according to the Dutch Medical Research Involving Human Subjects Act (WMO). All data were handled in line with the Helsinki Declaration.

Data selection

From the LROI register, we extracted data of patients aged between 50 and 80 years, who had primary TKA in the Netherlands in the period 2014-2019. The OAI cohort contained patients aged between 45 and 79 years. The OAI website provides detailed information about exclusion criteria and dropout rates prior to the enrollment visit. Data of patients with primary TKA was only extracted when patients were diagnosed with osteoarthritis.

Data collection

From the LROI register, patient and procedure characteristics as well as PROMs were extracted. Patient characteristics contained: gender, age, diagnosis, body mass index (BMI), smoking, operation side, year of operation and previous surgery on the affected joint. Procedure characteristics contained: type of surgery, fixation, and surgical approach. The extracted PROM scores were the Numeric Rating Scale (NRS) for pain at rest and during activity rated on a 0-10 rating scale, the Oxford Knee Score (OKS)(0-48)[11] and the Knee injury and Osteoarthritis Outcome Score-Physical Function Short-form (KOOS-PS)(0-100).[12] Patients were asked about satisfaction with their TKA through the following question: “How satisfied are you (in general) about the results of your knee surgery?” which was rated on an NRS from 0 to 10. In addition, two anchor questions on daily pain and functioning were asked: “How did your general daily pain/functioning change after the surgery on your knee?” (1 “very deteriorated”, 2 “deteriorated”, 3 “a little deteriorated”, 4 “unchanged”, 5 “a little improved”, 6 “improved”, or 7 “very improved”).

The OAI collected a large set of PROM scores at baseline and during each annual follow-up visit. We extracted baseline descriptive data (age and sex) and the following PROM scores: patient global assessment of knee impact rated on a 0-10 rating scale, global knee pain severity (not activity-specific) during the past 7 days rated on a 0-10 rating scale, the Knee injury and Osteoarthritis Outcome Score (KOOS)(0-100)[13], and the Western Ontario and McMasters Osteoarthritis index (WOMAC) Pain (0-20), Physical Function (0-68) and Stiffness (0-8) Likert scales.[14] Scores on these Likert scales are summed for the WOMAC total score (0-96). TKA was a time-varying treatment in this cohort as TKA interventions occurred asynchronously during follow-up. Therefore, the time between the annual follow-up visits and TKA surgery date varied for each patient. For the selection of preoperative PROMs, we used the visit prior to surgery (with a maximum of 1.5 year preoperative) and for the postoperative PROMs the closest visit after surgery (within the window of 1 to 2.5 years postoperative).

Statistical Analyses

A list of dichotomous definitions was used (Table 1), partly derived from an inventory review[6], and partly composed by an Expert Advisory Group within a qualitative study (submitted). Only dichotomous definitions tested at 12 months follow-up, or measuring a change score from preoperative to 12 month postoperative were added to the list. The following performance aspects of definitions were examined: the prevalence of poor responders, overlap of definitions and discriminative accuracy, measured with sensitivity, specificity, positive predictive value (PPV), negative predicted value (NPV) and Youden index. The proportion of overlap for all pairs of definitions within each dataset was calculated by Cohen’s kappa with 95% confidence intervals. Kappa values of 0.0 - 0.20 were considered “slight”, 0.21 - 0.4 “fair”, 0.41 - 0.60 “moderate”, 0.61 - 0.80 “substantial”, and 0.81 - 1.00 “almost perfect”.[15] Lack of overlap indicates that different outcome domains are being measured whereas almost perfect overlap of poor responders might indicate that different definitions might be interchangeable. Currently, there is no gold standard for measuring poor response to TKA. Therefore, the NRS satisfaction (0-10) in the LROI dataset was dichotomized (>6 ‘responder’ vs ≤6 ‘poor responder’) and used as anchor to calculate the discriminative accuracy of all examined definitions of poor response. In order to ensure robustness of the NRS satisfaction anchor, other outcome domains (e.g. pain and physical functioning) and different NRS cut-off values were tested as anchors. The results of different anchors were discussed and the final anchor was established on the basis of expert opinion. For the OAI, the measure Global assessment of Knee Impact (0-10) was dichotomized (<4 ‘responder’ vs ≥4 ‘poor responder’) for the analysis of discriminative accuracy. In addition, 95% confidence intervals (CI) for sensitivity and specificity, PPV and NPV and the Youden index (J = sensitivity + specificity – 1) were calculated. Statistical analyses were performed within the separate datasets and these were not combined or compared. All statistical analyses were performed in STATA version 17.0 (StataCorp, College Station, Texas).

Table 1. Definitions and prevalence of poor responders

Definition	Prevalence of poor responders %	
	LROI dataset	OAI dataset
Dichotomous definitions derived from an inventory review		
NRS pain >40 (0-100) [16,17]	12.6%	19.6%
OKS pain & functioning absolute improvement ≤5 [18]	13.1%	x
OKS pain & functioning absolute improvement ≤6 [19]	15.2%	x
NRS functioning >40 (0-100)*,** [20]	x	5.9%
OMERACT-OARSI responder criteria (WOMAC pain & functioning and global score): Non-responder: (<50% improvement and <20 absolute change in either pain or function) OR (no improvement in at least 2 of the 3 following: <20% improvement and <10 absolute change in either pain, function or patient's global assessment) [21-23]	x	27.5%
WOMAC pain, stiffness & functioning absolute improvement <7.5 [18]	x	18.1%
Dichotomous definitions composed by an Expert Advisory Group within a qualitative study		
<50% improvement and <20 absolute change in pain	24.7%	17.4%
<30% improvement and <10 absolute change in knee functioning	20.7%	23.0%
(<50% improvement and <20 absolute change in pain) OR (<30% improvement and <10 absolute change in knee functioning)	34.7%	27.5%
No improvement on transition question on change in pain***	10.1%	x
No improvement on transition question on change in daily knee functioning***	11.6%	x
No improvement on transition question on change in pain OR daily knee functioning	14.9%	x
(<50% improvement and <20 absolute change in pain) OR (No improvement on transition question on change in pain or daily knee functioning)	30.8%	x
(<30% improvement and <10 absolute change in knee functioning) OR (No improvement on transition question on change in pain or daily knee functioning)	27.9%	x
(<20% improvement and <10 absolute change in pain or knee functioning) AND (No improvement on transition question on change in pain or daily knee functioning)	14.9%	x

Abbreviations: NRS, Numeric Rating Scale; OKS, Oxford Knee Score; OMERACT-OARSI, Outcome Measures in Arthritis Clinical Trials-Osteoarthritis Research Society International; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index;
*NRS functioning derived from the original definition (International Knee Society score (IKSS) functioning <60);
**WOMAC functioning scale was used for the NRS functioning in the OAI dataset;
***Transition questions on change in pain and daily knee functioning after TKA range of 1 to 7, with 1 representing very deteriorated and 7 representing very improved. A score <4 was categorized as 'poor response'.

Results

Data characteristics

In our datasets, a total of 114,092 patients in the LROI and 416 patients (63 at baseline and 353 during follow-up) in the OAI cohort had at least one primary TKA (Figure 1). In the OAI cohort, 33 patients were excluded as they received a partial TKA, and another 34 patients were excluded as they were not diagnosed with OA. 99,531 and 255 patients (LROI and OAI, respectively) had undergone unilateral TKA and 14,561 and 98 patients (LROI and OAI, respectively) had bilateral TKA. Patients with bilateral TKA were excluded from our dataset as scores might be influenced by complaints on the left and right knee. Another 79,190 patients were excluded from our LROI dataset as they had incomplete preoperative and 12-month postoperative PROM questionnaires. 8,066 and 51 patients (LROI and OAI respectively) had missing data on PROM scores that were used for the definitions or as anchor and were excluded. A total of 12,275 (LROI) and 204 (OAI) patients with primary TKA were included in our analysis. The preoperative patient characteristics of the LROI and OAI datasets are reported in Table 2.

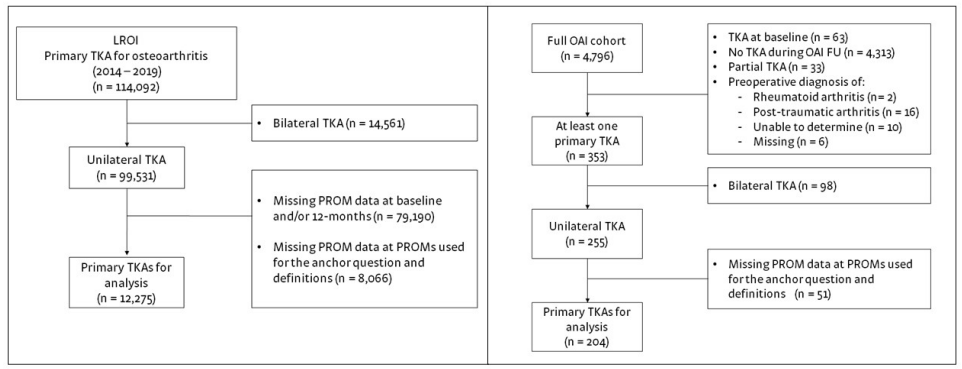


Figure 1. Flowchart showing data selection based on the LROI register and OAI database

Table 2. Preoperative patient characteristics and clinical details

Variable	LROI	OAI
	12,275	204
Female, *n* (%)	7,253 (59%)	125 (61%)
Age, mean years (SD)	67.5 (7.2)	64.6 (8.4)
BMI, mean (SD)	29.5 (4.8)	29.8 (4.4)
Baseline pain measured on a NRS (0 (no pain) - 10 (severe pain)), mean (SD)	5.3 (2.5)	6.4 (2.5)

Abbreviations: BMI, body mass index; NRS, numeric rating scale; SD, standard deviation

Anchor Questions

In the LROI dataset, the NRS satisfaction (0-10) was used as anchor, 2,335 patients (19%) were defined as poor responders (defined as ≤6) and 9,940 patients (81%) as responders (defined as >6) (Figure 2). The global assessment of knee impact score (0-10) was used as anchor in the OAI dataset, 30 patients (15%) were defined as poor responders (defined as ≥4) and 174 patients (85%) as responders (defined as <4).

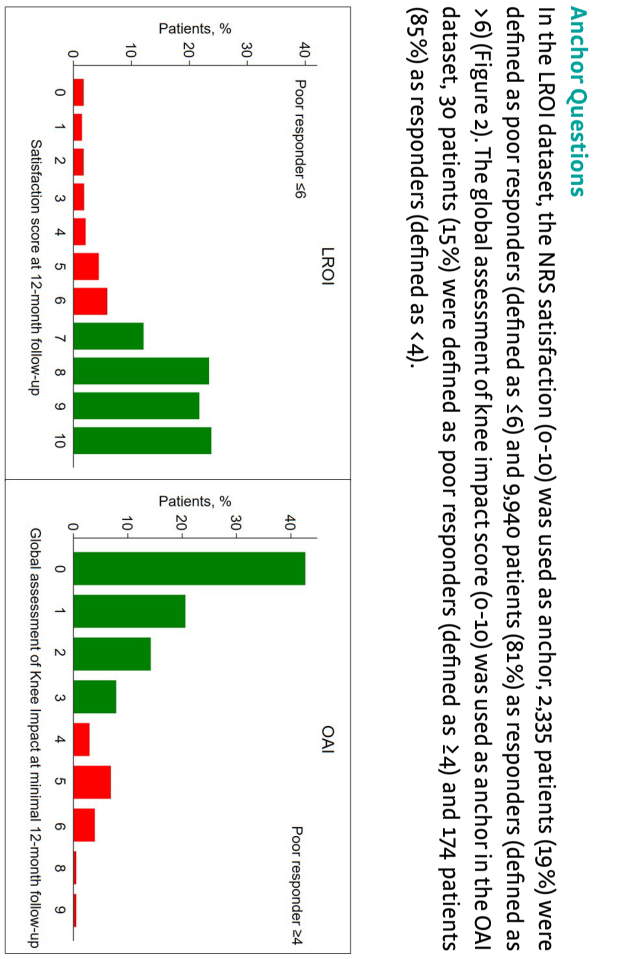


Figure 2. Distribution of answering options on anchor questions in the LROI and OAI datasets

Prevalence of poor responders

Table 1 shows the prevalence of poor responders of examined definitions within each dataset. The median (25^{th} – 75^{th} percentile) prevalence of poor responders was 18.1% (14.0% - 26.1%).

Overlap of definitions

The median (25^{th} – 75^{th} percentile) Cohen's kappa was 0.41 (0.32 - 0.59). A matrix overview of the Cohen's kappa scores is shown in Supplementary Table 1.

Discriminative accuracy

Table 3 shows the sensitivity, specificity, PPV and NPV with 95% CI and Youden's J-statistic of examined definitions of poor response to TKA within each dataset. The median (25^{th} – 75^{th} percentile) sensitivity was 0.47 (0.41 - 0.55), the median (25^{th} – 75^{th} percentile) specificity was 0.86 (0.82 - 0.94), the median (25^{th} – 75^{th} percentile) PPV was 0.45 (0.34 – 0.62), the median (25^{th} – 75^{th} percentile) NPV was 0.89 (0.87 – 0.89) and the median (25^{th} – 75^{th} percentile) Youden's J-statistic was 0.36 (0.24 - 0.45).

Table 3. Discriminative accuracy of definitions

	LROI dataset					OAI dataset				
Definition	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**J**	**PPV**	**NPV**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**J**	**PPV**	**NPV**
Dichotomous definitions derived from an inventory review										
NRS pain scale >40	0.39 (0.37–0.41)	0.93 (0.93–0.94)	0.32	0.58 (0.56–0.61)	0.86 (0.86–0.87)	0.60 (0.41–0.77)	0.87 (0.82–0.92)	0.47	0.45 (0.29–0.62)	0.93 (0.88–0.96)
OKS pain & functioning absolute improvement ≤5	0.43 (0.41–0.45)	0.94 (0.94–0.94)	0.37	0.63 (0.60–0.65)	0.88 (0.87–0.88)	x	x	x	x	x
OKS pain & functioning absolute improvement ≤6	0.48 (0.46–0.50)	0.93 (0.92–0.93)	0.41	0.60 (0.58–0.62)	0.88 (0.88–0.89)	x	x	x	x	x
NRS functioning >40	x	x	x	x	x	0.23 (0.10–0.42)	0.97 (0.93–0.99)	0.21	0.58 (0.28–0.85)	0.88 (0.83–0.92)
OMERACT-OARSI responder criteria (WOMAC pain & functioning and global score): Non-responder: (<50% improvement and <20 absolute change in either pain or function) OR (no improvement in at least 2 of the 3 following: <20% improvement and <10 absolute change in either pain, function or patient's global assessment)	x	x	x	x	x	0.43 (0.26–0.63)	0.75 (0.68–0.82)	0.19	0.23 (0.13–0.36)	0.89 (0.82–0.93)
WOMAC pain, stiffness & functioning absolute improvement <7.5	x	x	x	x	x	0.27 (0.12–0.46)	0.83 (0.77–0.89)	0.10	0.22 (0.10–0.38)	0.87 (0.81–0.92)
Dichotomous definitions composed by an Expert Advisory Group within a qualitative study										
<50% improvement and <20 absolute change in pain	0.57 (0.55–0.59)	0.83 (0.82–0.83)	0.39	0.44 (0.42–0.45)	0.89 (0.88–0.90)	0.25 (0.11–0.45)	0.84 (0.77–0.89)	0.09	0.21 (0.09–0.39)	0.87 (0.80–0.92)
<30% improvement and <10 absolute change in knee functioning	0.48 (0.46–0.50)	0.86 (0.85–0.86)	0.33	0.44 (0.42–0.46)	0.87 (0.87–0.88)	0.43 (0.26–0.63)	0.81 (0.74–0.86)	0.24	0.28 (0.16–0.43)	0.89 (0.83–0.94)
(<50% improvement and <20 absolute change in pain) OR (<30% improvement and <10 absolute change in knee functioning)	0.70 (0.68–0.72)	0.74 (0.73–0.75)	0.44	0.39 (0.37–0.40)	0.91 (0.91–0.92)	0.47 (0.28–0.66)	0.76 (0.69–0.82)	0.23	0.25 (0.14–0.38)	0.89 (0.83–0.94)
No improvement on transition question on change in pain	0.40 (0.38–0.42)	0.97 (0.97–0.97)	0.36	0.75 (0.72–0.77)	0.87 (0.87–0.88)	x	x	x	x	x
No improvement on transition question on change in daily knee functioning	0.41 (0.39–0.43)	0.95 (0.95–0.96)	0.36	0.67 (0.65–0.70)	0.87 (0.87–0.88)	x	x	x	x	x
No improvement on transition question on change in pain OR daily knee functioning	0.52 (0.50–0.54)	0.94 (0.93-0.94)	0.46	0.66 (0.64–0.68)	0.89 (0.89–0.90)	x	x	x	x	x

5

(<50% improvement and <20 absolute change in pain) OR (No improvement on transition question on change in pain or daily knee functioning)	0.72 (0.70 – 0.74)	0.79 (0.78 – 0.80)	0.51	0.44 (0.43 – 0.46)	0.92 (0.92 – 0.93)	×	×	×	×	×	×
(<30% improvement and <10 absolute change in knee functioning) OR (No improvement on transition question on change in pain or daily knee functioning)	0.68 (0.66 – 0.70)	0.82 (0.81 – 0.82)	0.50	0.47 (0.45 – 0.48)	0.92 (0.91 – 0.92)	×	×	×	×	×	×
(<20% improvement and <10 absolute change in pain or knee functioning) AND (No improvement on transition question on change in pain OR daily knee functioning)	0.52 (0.50 – 0.54)	0.94 (0.93 – 0.94)	0.46	0.66 (0.64 – 0.68)	0.89 (0.89 – 0.90)	×	×	×	×	×	×

Abbreviations: CI, confidence interval; J, Youden index; NPV, negative predictive value; PPV, positive predictive value

Discussion

To our knowledge, this is the first study that compared the performance of various definitions of poor response to TKA. In general, the prevalence of poor responders was discordant across the different definitions in this study. The majority of pairs of definitions showed a fair or moderate agreement suggesting that different definitions of poor response to TKA measure different underlying outcome domains. Definitions showed a relatively low sensitivity, PPV and Youden's J-statistic but moderate to good specificity and NPV suggesting that absence of poor response can be classified with confidence. Our findings provide insight in the extent to which definitions of poor response relate to each other, thus supporting knee specialists and researchers to interpret and compare results about response to TKA across studies.

In the literature, it is assumed that the proportion of patients with poor response to TKA is about 15 to 20%.[1,4,5] This percentage seems to be generally accepted in the absence of a gold standard. In this study, the prevalence of poor responders according to different definitions, ranging from 5.8% to 35.6%, scatters around a proportion of 15-20%, suggesting that the "true estimate" of poor responders lies indeed around 15-20%. However, some definitions yielded much lower (<10%) and/or much higher proportions of poor responders (>25%). Four definitions could be tested in both the LROI and OAI dataset and the average difference in prevalence values between both datasets was 6%, suggesting that definitions perform fairly constant in different datasets.

The highest Cohen's kappa values were measured for combinations of definitions that measured the same underlying outcome domain or that partially used the same criteria. Perhaps unsurprisingly, a perfect overlap (1.00) was measured for the definition *'No improvement on transition question on change in pain OR daily knee functioning'* versus the definitions *'(<20% improvement and <10 absolute change in pain rest or knee functioning) AND (No improvement on transition question on change in pain or daily knee functioning)'*. We observed lowest agreement between definitions with an absolute cut-off and definitions that contained both a relative and absolute change. The latter observation applies to combinations that measured the same underlying outcome domain as well as to combinations that did not measure the same underlying outcome domain.

Definitions that perform best in terms of discriminative accuracy consist of multiple outcome domains such as pain and functioning, and contain both a relative and absolute change from preoperative to 12 months postoperative, such as the Outcome Measures in Rheumatology (OMERACT) and Osteoarthritis Research Society International (OARSI) set of responder criteria.[24] The OMERACT-OARSI responder criteria, originally developed to examine the effect of pharmacological interventions in OA are commonly used in the literature to identify responders in OA.[24] Two studies used this OMERACT-OARSI responder set to examine the prevalence of non-response to TKA and found non-response rates of 15%, 19.3% and 24.4%.[21,22] These rates are slightly lower compared to a 27.5% OMERACT-OARSI non-response rate in our OAI dataset, which might be explained by a different patient cohort and different inclusion criteria used for the OAI dataset. Our findings concerning the discriminative accuracy of the OMERACT-OARSI criteria to identify poor response after TKA are comparable to a previous study on the accuracy of the OMERACT-OARSI criteria to identify patients who worsened despite 3 months of conservative treatment, demonstrating high specificity, but low

sensitivity and a low Youden index.[25] Remarkably, no definition (set of responder or worsening criteria) has a good Youden index score.[25] Hence, the OMERACT-OARSI responder criteria constitute a definition that can classify patients who do not experience a poor response, but is less accurate in classifying patients as poor responders.

Limitations

The main limitation of the present study are the chosen anchors for the assessment of discriminative accuracy in both datasets. In the absence of a true gold standard, the outcome domains, cut-off values and the timing of assessment were arbitrarily chosen. The choice for this type of gold standard introduces some problems with regard to the validity of the anchor.[26] A single domain (e.g. pain) is assumed to be less reliable, and thus less valid than an anchor that includes multiple domains (e.g. pain combined with daily knee functioning). A general domain was chosen as anchor based on results of a prior qualitative study that identified (dis)satisfaction as a central theme of poor response to TKA (submitted). Moreover, a general domain is a commonly used anchor in this type of studies.[27,28] The cut-off values were discussed and established on the basis of expert opinion. The timing of assessment, 12 months after TKA, was also based on the prior qualitative study where patients and knee specialists were asked to indicate the moment of follow-up for the assessment of the definition (submitted). In summary, our data on the discriminative accuracy results are determined by the choices made for the anchors. Riddle et al.[29] already reported not to rely on arbitrary cut-offs, and preferred a non-biased statistical model-based approach to categorize good versus poor outcome. However, a major limitation of these types of models is that dichotomization can only be determined afterwards and results are only generalizable after external validation. Next, with a response rate of 38% in matched pre- and 12 months postoperative PROMs in the LROI register there might be some random bias, because of the considerable variation in response rate among hospitals. From previous research is know that patients who do not complete the PROMs of the LROI register are in general slightly older and less fit.[30] Another limitation concerns the time varying data in the OAI dataset. We therefore set minimum and maximum time windows of assessment prior to and after the surgery date considering that the postoperative poor response rate may decline over time.[22]

In a parallel study by our research group we performed a Delphi exercise to assess the face validity and feasibility, and to prioritize definitions from an initial list of 34 definitions. A gold standard for measuring poor response to TKA is still lacking. Therefore, the generation of a prioritized list of definitions by international experts provides a useful resource for helping to guide clinicians and researchers to choose a definition that corresponds to their aim and setting of measuring poor response to TKA. Future research into the patient prioritization of outcomes of poor response to TKA is necessary to identify the most relevant definition of poor response to TKA. With this study, international experts will be informed about the results and encouraged to choose a definition that can be used to measure poor response of TKA worldwide, making it possible to compare the prevalence over time and across hospitals and countries.

Conclusions

In conclusion, the wide range in prevalence of poor response using different definitions and the lack of overlap stress the need for an unambiguous definition of poor response to TKA. None of the definitions we examined adequately classified poor responders of TKA. In contrast, absence of poor response could be classified with confidence. A future definition of poor response to TKA should preferably include multiple outcome domains and combine a relative and absolute change.

Acknowledgements

We would like to thank the members of the Expert Group (apart from the authors: S. van Onsem, A.J. Porteous, O. Rolfson and J.A. Singh) for their contributions to the discussion and establishment on the cut-off values for the anchor questions.

The OAI is a public-private partnership comprised of five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2-2262) funded by the National institutes of Health, a branch of the Department of Health and Human Services, and conducted by the OAI Study Investigators. Private funding partners include Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer, Inc. Private sector funding for the OAI is managed by the Foundation for the National Institutes of Health. This paper was prepared using an OAI public use data set and does not necessarily reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners.

5

5

References

1. Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. BMJ open 2012;2(1):1-12.
2. Gunaratne R, Pratt DN, Banda J, Fick DP, Khan RJK, Robertson BW. Patient Dissatisfaction Following Total Knee Arthroplasty: A Systematic Review of the Literature. J Arthroplasty 2017;32(12):3854-3860.
3. Canovas F, Dagneaux L. Quality of life after total knee arthroplasty. Orthop Traumatol Surg Res 2018;104(1S):41-46.
4. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010;468(1):57-63.
5. Halawi MJ, Jongbloed W, Baron S, Savoy L, Williams VJ, Cote MP. Patient Dissatisfaction After Primary Total Joint Arthroplasty: The Patient Perspective. J Arthroplasty 2019;34(6):1093-1096.
6. Te Molder MEM, Smolders JMH, Heesterbeek PJC, van den Ende CHM. Definitions of poor outcome after total knee arthroplasty: an inventory review. BMC Musculoskelet Disord 2020;21(1):378.
7. Price AJ, Longino D, Rees J, et al. Are pain and function better measures of outcome than revision rates after TKR in the younger patient? Knee 2010;17(3):196-9.
8. (LROI) DAR. Completeness (proportion [%] per joint) of the registration of procedures in the LROI in 2013-2020. 2021. Retrieved from: https://www.lroi-rapportage.nl/data-quality-coverage-and-completeness19. Date last accessed: 12 April 2022.
9. (LROI) DAR. Pre-operative, 6 months and 12 months postoperative response percentage of patients who underwent a TKA for osteoarthritis per preoperative PROMs registering hospital (n=92) in the Netherlands in 2014-2020. 2021. Retrieved from: https://www.lroi-report.nl/knee/proms/response/. Date last accessed: 12 April 2022.
10. Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis Cartilage 2008;16(12):1433-41.
11. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. The Journal of bone and joint surgery British volume 1998;80(1):63-9.
12. Perruccio AV, Stefan Lohmander L, Canizares M, et al. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS) - an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008;16(5):542-50.
13. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. The Journal of orthopaedic and sports physical therapy 1998;28(2):88-96.
14. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15(12):1833-40.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-74.
16. Brander V, Gondek S, Martin E, Stulberg SD. Pain and depression influence outcome 5 years after knee replacement surgery. Clin Orthop Relat Res 2007;464:21-6.
17. Brander VA, Stulberg SD, Adams AD, et al. Predicting total knee replacement pain: a prospective, observational study. Clin Orthop Relat Res 2003;(416):27-36.
18. Alzahrani K, Gandhi R, Debeer J, Petruccelli D, Mahomed N. Prevalence of clinically significant improvement following total knee replacement. J Rheumatol 2011;38(4):753-9.
19. Filbay SR, Judge A, Delmestri A, Arden NK. Evaluating Patients' Expectations From a Novel Patient-Centered Perspective Predicts Knee Arthroplasty Outcome. J Arthroplasty 2018;33(7):2146-2152.e4.
20. Dowsey MM, Nikpour M, Dieppe P, Choong PF. Associations between pre-operative radiographic changes and outcomes after total knee joint replacement for osteoarthritis. Osteoarthritis Cartilage 2012;20(10):1095-102.
21. Dowsey MM, Spelman T, Choong PF. Development of a Prognostic Nomogram for Predicting the Probability of Nonresponse to Total Knee Arthroplasty 1 Year After Surgery. J Arthroplasty 2016;31(8):1654-60.
22. Riddle DL, Golladay GJ, Jiranek WA, Perera RA. External Validation of a Prognostic Model for Predicting Nonresponse Following Knee Arthroplasty. J Arthroplasty 2017;32(4):1153-1158.
23. Weber M, Craiovan B, Woerner ML, Schwarz T, Grifka J, Renkawitz TF. Predictors of Outcome After Primary Total Joint Replacement. J Arthroplasty 2018;33(2):431-435.
24. Pham T, van der Heijde D, Altman RD, et al. OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. Osteoarthritis Cartilage 2004;12(5):389-99.
25. Mahler E, den Broeder AA, Woodworth TG, et al. How should worsening in osteoarthritis be defined? Development and initial validation of preliminary criteria for clinical worsening in knee and hip osteoarthritis. Scand J Rheumatol 2017;46(5):396-406.
26. Terwee CB, Roorda LD, Dekker J, et al. Mind the MIC: large variation among populations and methods. J Clin Epidemiol 2010;63(5):524-34.
27. Beard DJ, Harris K, Dawson J, et al. Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. J Clin Epidemiol 2015;68(1):73-9.
28. Most J, Hoelen TA, Spekenbrink-Spooren A, Schotanus MGM, Boonen B. Defining Clinically Meaningful Thresholds for Patient-Reported Outcomes in Knee Arthroplasty. J Arthroplasty 2022;37(5):837-844.
29. Riddle DL, Dumenci L. Classifications of good versus poor outcome following knee arthroplasty should not be defined using arbitrary criteria. BMC Musculoskelet Disord 2020;21(1):604.
30. (LROI) DAR. Representativeness of the results of the PROMs questionnaire total hip arthroplasty. 2017. Retrieved from: https://issuu.com/novlroi/docs/lroi_magazine_2017. Date last accessed: 2 May 2023.

Supplementary table 1. Results of overlap of definitions based on the LROI and OAI datasets

LROI / OAI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. NRS pain scale >40		0.35 (0.33–0.37)	0.36 (0.34–0.38)				0.60 (0.58–0.61)	0.25 (0.23–0.27)	0.42 (0.41–0.44)	0.42 (0.40–0.45)	0.33 (0.30–0.35)	0.37 (0.35–0.39)	0.49 (0.47–0.50)	0.29 (0.27–0.31)	0.37 (0.35–0.40)
2. OKS pain & functioning absolute improvement ≤5			0.91 (0.90–0.92)				0.35 (0.33–0.37)	0.45 (0.43–0.47)	0.34 (0.33–0.36)	0.43 (0.41–0.45)	0.39 (0.36–0.41)	0.41 (0.39–0.44)	0.34 (0.32–0.35)	0.42 (0.40–0.44)	0.41 (0.39–0.44)
3. OKS pain & functioning absolute improvement ≤6							0.38 (0.36–0.39)	0.46 (0.44–0.48)	0.37 (0.36–0.39)	0.43 (0.40–0.45)	0.38 (0.36–0.40)	0.42 (0.40–0.44)	0.37 (0.35–0.39)	0.45 (0.43–0.46)	0.42 (0.40–0.44)
4. NRS functioning >40*	0.32 (0.16–0.49)														
5. Non-responder on OMERACT-OARSI responder criteria	0.14 (-0.01–0.28)			0.12 (0.00–0.24)											
6. WOMAC pain, stiffness & functioning absolute improvement <7.5	0.21 (0.06–0.37)			0.31 (0.14–0.48)	0.60 (0.47–0.73)										
7. <50% improvement and <20 absolute change in pain	0.29 (0.12–0.46)			0.12 (-0.04–0.27)	0.44 (0.30–0.58)	0.45 (0.29–0.61)		0.32 (0.30–0.34)	0.77 (0.75–0.78)	0.32 (0.30–0.34)	0.26 (0.24–0.28)	0.32 (0.31–0.34)	0.85 (0.84–0.86)	0.37 (0.35–0.39)	0.33 (0.31–0.34)
8. <30% improvement and <10 absolute change in knee functioning	0.20 (0.05–0.35)			0.23 (0.09–0.38)	0.65 (0.53–0.77)	0.82 (0.72–0.92)	0.51 (0.36–0.66)		0.66 (0.64–0.67)	0.27 (0.25–0.29)	0.28 (0.26–0.30)	0.32 (0.30–0.34)	0.34 (0.32–0.35)	0.81 (0.80–0.82)	0.32 (0.30–0.34)
9. (<50% improvement and <20 absolute change in pain) OR (<30% improvement and <10 absolute change in knee functioning)	0.27 (0.12–0.42)			0.19 (0.06–0.31)	0.61 (0.48–0.73)	0.71 (0.60–0.82)	0.68 (0.56–0.80)	0.88 (0.81–0.96)		0.26 (0.24–0.27)	0.24 (0.22–0.26)	0.30 (0.28–0.32)	0.74 (0.72–0.75)	0.66 (0.65–0.68)	0.30 (0.28–0.32)
10. No improvement on transition question on change in pain**											0.57 (0.55–0.60)	0.78 (0.76–0.80)	0.40 (0.39–0.42)	0.45 (0.43–0.47)	0.78 (0.76–0.80)
11. No improvement on transition question on change daily knee functioning**												0.85 (0.84–0.87)	0.46 (0.44–0.47)	0.51 (0.49–0.52)	0.85 (0.84–0.87)
12. No improvement on transition question on change in pain OR daily knee functioning													0.57 (0.55–0.58)	0.63 (0.61–0.64)	1.00 (0.99–1.00)
13. (<50% improvement and <20 absolute change in pain) OR (No improvement on transition question on change in pain or daily knee functioning)														0.55 (0.54–0.57)	0.57 (0.55–0.58)
14. (<30% improvement and <10 absolute change in knee functioning) OR (No improvement on transition question on change in pain or daily knee functioning)															0.62 (0.61–0.64)
15. (<20% improvement and <10 absolute change in pain rest or knee functioning) AND (No improvement on transition question on change in pain or daily knee functioning)															
Slight overlap: 0–0.20			Fair overlap: 0.21–0.40			Moderate overlap: 0.41–0.60			Substantial overlap: 0.61–0.80			almost perfect overlap: ≥ 0.81			

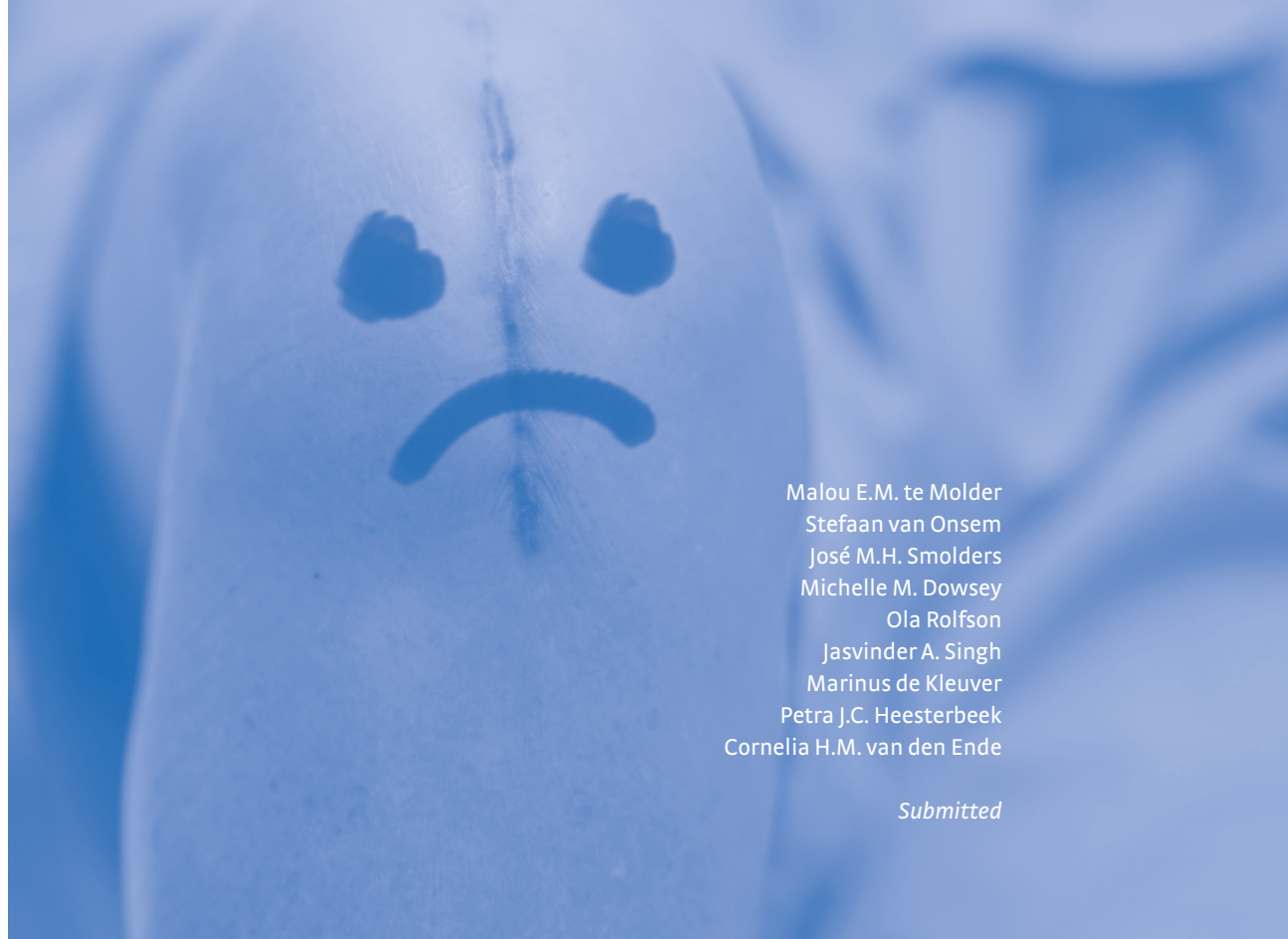
Abbreviations: NRS, Numeric Rating Scale; OKS, Oxford Knee Score; OMERACT-OARSI, Outcome Measures in Arthritis Clinical Trials-Osteoarthritis Research Society International; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index

*WOMAC functioning scale was used for the NRS functioning in the OAI dataset;

**Transition questions on change in pain and daily knee functioning after TKA range of 1 to 7, with 1representing verydeteriorated and 7 representing very improved. A score <4 was categorized as 'poor response'.

Chapter 6

International consensus-based ranking of definitions for poor response to primary total knee arthroplasty: A Delphi study



Malou E.M. te Molder
Stefaan van Onsem
José M.H. Smolders
Michelle M. Dowsey
Ola Rolfson
Jasvinder A. Singh
Marinus de Kleuver
Petra J.C. Heesterbeek
Cornelia H.M. van den Ende

Submitted

Abstract

Introduction
This study aimed to rank definitions for measuring poor response one year after TKA, after assessing the face validity and feasibility of existing or newly proposed definitions.

Materials and Methods
An international, three-round, online modified Delphi study was conducted with sixty-nine panelists from twenty-three countries. Definitions were derived from a literature review or were newly proposed by an expert group. Panelists rated the face validity and feasibility of definitions, and could propose additional new definitions in round 1. Panelists reconsidered their rating of existing definitions, and rated newly suggested definitions (round 2). Definitions with a median score for face validity <6.5 were removed from the list, and panelists distributed 100 points among the remaining definitions for ranking (round 3).

Results
Fifty-one panelists completed all three rounds (response rate 74%), and the prioritized list of definitions in round 3 comprised seventeen definitions. A definition on (dis)satisfaction with the outcome of TKA obtained the highest scores for face validity and feasibility (7.5, and 8.5 respectively), and a definition reflecting change since pre-operative status from the perception of patients (i.e., their opinion on change in pain, daily knee functioning) was highest prioritized. In general, definitions reflecting change from the perception of patients were higher prioritized than definitions requiring both preoperative and postoperative assessment of validated questionnaires.

Conclusions
This study identified seventeen potential definitions of poor response to TKA, of which six we believe justify further study and potential implementation in quality assessment studies. Remarkably, all six were patient-centered and none were clinician-centered.

Keywords
Total knee arthroplasty, Definitions, Poor response, Delphi study

Introduction

Over the last decades total knee arthroplasty (TKA) has been widely established as a successful and cost-effective procedure for advanced symptomatic osteoarthritis (OA). Worldwide an average of around 135 TKAs per 100,000 persons are performed[1,2], and utilization is projected to increase continuously over the next decade.[3] The massive growth in demand for TKA will inevitably place a major burden on the future healthcare systems across the world, with increasing costs and limited resources. However, despite improvement in knee pain and disability, 10 to 20% of patients report that they are dissatisfied due to insufficient pain relief, limitations in physical functioning and/or unmet expectations.[4-6]

To be able to improve these dissatisfaction rates, a definition on poor outcome after TKA is needed, to allow for an actionable quality improvement cycle. A variety of dichotomous definitions of poor response to TKA comprising one or more different dimensions of outcome have been described in the literature[7], to quantify the proportion of patients with poor response to TKA. This large variety of definitions impedes the comparisons of poor response to TKA over time and across hospitals and countries. The need for a multidimensional combination of outcome domains (e.g. pain and function) has been recognized to describe failure (i.e. poor response) after TKA[8], but to date, an international accepted definition with good performance for measuring poor response to TKA is lacking (submitted).

Previous research has focused on identifying a core outcome domain set for total joint arthroplasty (TJA) clinical trials and meaningful improvements for the domain measures.[9-11] So far, few studies addressed outcome domains that are important for measuring failure after TKA. One study focusing on defining TJA failure from the patient's perspective identified refractory pain after TJA as more important than surgical failure (i.e. complications, revisions).[12] Recently, our research group conducted a qualitative study (submitted) among patients and knee specialists to explore consequences that might contribute to the negative appraisal of a TKA procedure. This study revealed new insights, including difficulties in the process of adapting to a TKA over time and the continuous annoying awareness of the artificial knee (submitted).

A definition of poor response to TKA after one year should encompass criteria to identify patients with an unfavorable course. This definition should specify the domain(s) or outcome measures (e.g., revision rate or physical functioning) and include criteria regarding the magnitude and nature of change (either relative or absolute change compared to preoperative status (transition question)) or a postoperative threshold beyond which the patient is considered to have a poor response. Furthermore, worldwide implementation of definitions requires that such a definition is both valid (the degree to which the definition is an adequate reflection of 'poor response one year after TKA') and feasible (the degree to which the definition is easy to use and assess worldwide). Thus, the definition should have good performance of discriminative accuracy, and should be feasible from an international point of view. The primary aim of this study was to seek consensus among international orthopedic knee experts regarding the face validity and feasibility of existing and newly proposed definitions for defining poor response one year after primary TKA. Secondary aim was to rank definitions to gain insight into the most important definitions for measuring poor response to TKA that require further exploration.

Material and methods

This three-round online modified Delphi study is reported in line with recommendations for the Conducting and REporting of DElphi Studies (CREDES)[13,14] and proposed Delphi study quality indicators.[15]

Project team and Expert Advisory Group

A project team was formed to conduct the study comprising two orthopedic knee surgeons from the Netherlands and Belgium (JS, SvO), two researchers with background in rheumatology and orthopedics (CvdE, PH) and a PhD student (MtM). An expert advisory group, involving the five project team members and four international key experts with expertise and scientific publications on defining outcome after TKA, was established. The four key experts included a professor, orthopedic surgeon from Sweden (OR), a professor, rheumatologist from the USA (JS), a professor, epidemiologist and nurse from Australia (MD) and a leading orthopedic surgeon and orthopedic engineer from the UK (AP). Members of the expert advisory group were not members of the Delphi panel.

Advisory Group meetings

The project team and expert advisory group met four times during the study (Figure 1). Two meetings were held prior to the first Delphi round, one meeting between Delphi round 2 and 3 and the last meeting was held after completion of round 3. In preparation for the first meeting, the members completed a survey hosted using SurveyMonkey[16] and administered via e-mail. In this survey, members were asked to rank domains of poor response originating from a prior (inventory) literature review[7] and results of a qualitative study (submitted). Furthermore, members were asked about their opinion and preference for the type of threshold (e.g., absolute cut-off value, absolute change, relative change) for definitions identifying poor response one year after TKA. During the preparatory meetings, members of the advisory group were informed about the results of the studies preceding this Delphi study[7] (submitted). The initial list of definitions for round 1 was created by combining existing definitions that emerged from a (inventory) literature review with definitions composed by the expert advisory group, based on discussion of the results of a qualitative study among patients and knee specialists, which explored consequences that may contribute to the negative appraisal of a TKA procedure (submitted). The list of definitions used in this study focused on assessment of response at one year after TKA as this time point was identified as the optimal timepoint for the assessment by patients and knee specialists (submitted), and because it is an important time point during routine follow-up. In addition to the definitions that emerged from the (inventory) literature review, new definitions were composed by the members, based on discussions of the results of the qualitative study (submitted). During the third meeting, adjustments of definitions were discussed based on comments that arose from the first two Delphi rounds. Furthermore, a threshold for removing definitions from the list that served as input for Delphi round 3 was discussed and established. Different thresholds were tested to find an optimal balance between the number of definitions to be ranked and the relevance of definitions. One team member (MtM) tested the following 3 different thresholds predefined by the expert advisory group: on the median score of face validity, a threshold median value of 6.5 and 7, and the value corresponding with the 25th percentile. The expert advisory group unanimously agreed on a median score lower than 6.5 as the threshold for removal. The final meeting was organized to discuss the results of the Delphi exercise and to formulate a list of prioritized definitions for defining poor response one year after TKA.

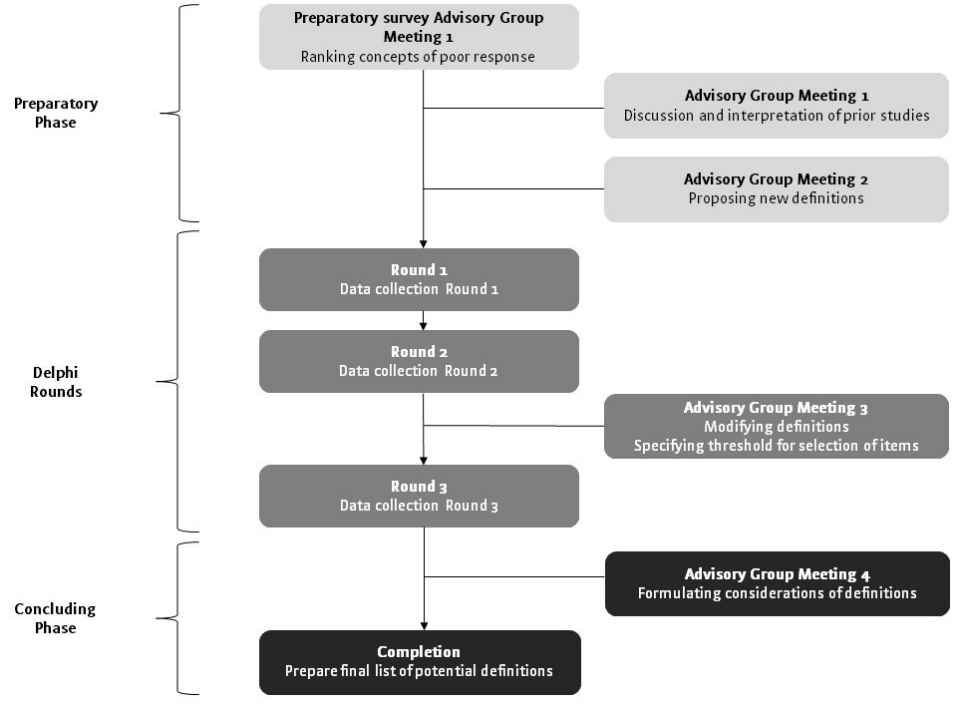


Figure 1. Flowchart of the Delphi process

Expert panel

There are no established guidelines on the optimal Delphi study panel size.[17] Therefore, a target of 50 panelists from at least 5 different countries worldwide, was set to ensure that international key stakeholders were sufficiently represented. Panelists were invited based on their recognized knowledge of the topic, their willingness to participate, and their intention to commit to the process.

Members of the expert advisory group were encouraged to share study details with other experts in the field and to recruit potential panelists. Panelists were invited via e-mail to participate and were asked to nominate additional potential panelists (snowball-sampling).[18] To ensure clinical and research expertise, they were included in the panel when they met the following eligibility criteria: (1) professional background as an orthopedic knee surgeon or orthopedic researcher; (2) (co)authored at least two publications on the outcome of TKA and/or performed at least 30 knee arthroplasties yearly; (3) able to communicate in English and use/access to internet and e-mail; (4) ability and willingness to respond to each Delphi survey within two weeks.

Delphi procedure

The panelists were emailed explaining the purpose, rationale, and the content of the study. To increase participation the panelists were asked to reply if they were willing to join and whether they intended to commit to the procedure. The procedure was performed between April and August 2021. It was decided a priori to include three rounds to increase convergence whilst minimizing participant attrition[19] (Figure 1). All three surveys were hosted using SurveyMonkey[16] and administered via e-mail. Reminders were provided via e-mail to help maximize response rates. All panelists who completed round 1 were subsequently emailed links to round 2 and 3. In each round the purpose and procedure of the current Delphi round were explained. Panelists remained anonymous and unknown to each other throughout the entire process.

Data collection

Round 1

The initial draft list of definitions was provided to the Delphi panel (supplementary table 1). Panelists were asked to score the face validity (the degree to which the definition is an adequate reflection of 'poor response one year after TKA') and feasibility (the degree to which the definition is easy to use and assess worldwide) of each definition on a scale of 0 (very low face validity or not feasible) to 10 (very high face validity or feasible) and to motivate each score. Free-text options were included at the end of the survey to allow panelists to suggest new definitions of poor response to TKA. Prior to round 1, the survey was pilot-tested by two independent researchers. This led to minor wording/structural changes for clarity.

Round 2

Each definition from round 1 was accompanied by a table showing the n (number of panelists that provided a score in round 1), minimum, maximum, mean and median score for that definition on face validity and feasibility in round 1 (illustrated in Figure 2). A summary of the face validity and feasibility comments were shown below the table.

Round 2 also included suggested definitions generated from the free-text responses of round 1. Panelists were asked to reconsider their rating of definitions from round 1 and to score the face validity and feasibility of the newly suggested definitions from round 1 (on a scale of 0 to 10). In round 2, free-text options for comments were included after each definition.

Definition 1 out of 25:

Numeric pain rating scale (scale 0-100) >40 at 1 year follow-up (yes/no)

Poor responder = if the answer is yes

Results	Face validity	Feasibility
N	69	69
Min	2	2
Max	10	10
Mean	6.5	8.1
median	7	9

Summary of the fact validity comments:

- pain can be influenced by other factors (depression, anxiety)
- pain is subjective and difficult to measure
- asking patients for pain is mandatory in every follow-up
- pain scale should be low because pain is the main reason for TKA
- pain is very important to patients
- persistent pain is the most significant indicator for poor response
- pain can not be the only parameter to identify poor outcome
- threshold is too high

Summary of the feasibility comments:

- Easy to use, to understand and to interpret for the patient
- Well known method
- Not time consuming
- NRS 0-10 is simpler and easier to evaluate than 0-100
- Prospectively measurable
- Very individual, subjective parameter, only potential is evaluating the trend before vs after TKA
- A good score but only for pain

face validity:

V

feasibility:

V

Figure 2. Round 2 survey item example

Round 3

In round 3, definitions with a median score for face validity lower than 6.5 were removed from the list of definitions. Panelists were asked to distribute 100 points over the remaining definitions (n=17) to rank the definitions of poor response. No free-text options were included in round 3 to minimize panelists' burden.

Data analysis

Qualitative analysis

The suggested definitions in round 1 and the comments motivating the scores for face validity and feasibility in round 1 and 2 were collected in Microsoft Excel 365. Comments for face validity and feasibility scores were summarized per definition and duplicate suggested definitions in round 1 were removed by one project team member (MtM). The list with potential definitions (round 1) and summarized comments (round 1 and 2) was discussed and verified with one additional team member (CvdE). These comments were used to adapt draft versions of the remaining definitions for round 3. The list with adjusted definitions was discussed with the expert advisory group and further modified. A final list with potential definitions, was input for the third Delphi round.

Quantitative analysis

The mean (SD) face validity and feasibility scores, the sum score of face validity plus feasibility (mean face validity score plus mean feasibility score), the ranking of definitions (total ranking points per definition) and the percentage of panelists that scored at least 1 point for a definition were analyzed descriptively using Microsoft Excel 365 and STATA 13.0 (StataCorp, College Station, TX).

Results

Panelists response

105 potential panelists were nominated and screened, of whom 100 met the eligibility criteria (Figure 3). A total of 100 eligible panelists were invited to participate in the Delphi exercise, of whom 8 did not respond to the invitation, 1 was not willing to participate and 4 were not invited further to avoid one country being over-represented (13 Dutch panelists). The remaining 87 panelists were emailed the link to round 1. A total of 69 panelists completed round 1 and formed the Delphi panel. Rounds 2 and 3 were completed by 63 (91%) and 51 (74%) panelists, respectively. Reasons for non-response of the different rounds were not available. Table 1 provides the panelists' characteristics of participants that completed the first round.

Table 1. Characteristics of 69 panelists who completed Round 1 and formed the Delphi panel

	Number of panelists (%) (n = 69)
Women, (N(%))	6(8.7%)
Missing	1(1.4%)
Country of residency, (N(%))	
Netherlands	13(18.8%)
UK	9(13.0%)
USA	7(10.1%)
Australia	6(8.7%)
Italy	4(5.8%)
South Africa	3(4.3%)
Germany	3(4.0%)
Belgium	3(4.3%)
France	3(4.0%)
Norway	2(2.9%)
Austria	2(2.9%)
Indonesia	2(2.9%)
Denmark	2(2.9%)
Other [a]	11(15.9%)
Missing	1(1.4%)
Current professional role, (N(%))	
Clinician	39(56.5%)
Clinician & researcher	21(30.4%)
Researcher	8(11.6%)
Missing	1(1.4%)

[a] Other include the following countries: Finland, Slovenia, Indonesia, Scotland, Switzerland, Canada, Sweden, New Zealand, Greece, Spain and India

Round 1

Face validity and feasibility scores of the initial 25 included definitions are shown in supplementary table 2. 25 Panelists proposed 29 new different definitions of which 9 were added to round 2 on the basis of consensus among the members of the expert advisory group (supplementary table 3). Panelists took on average 19 min time to complete Round 1.

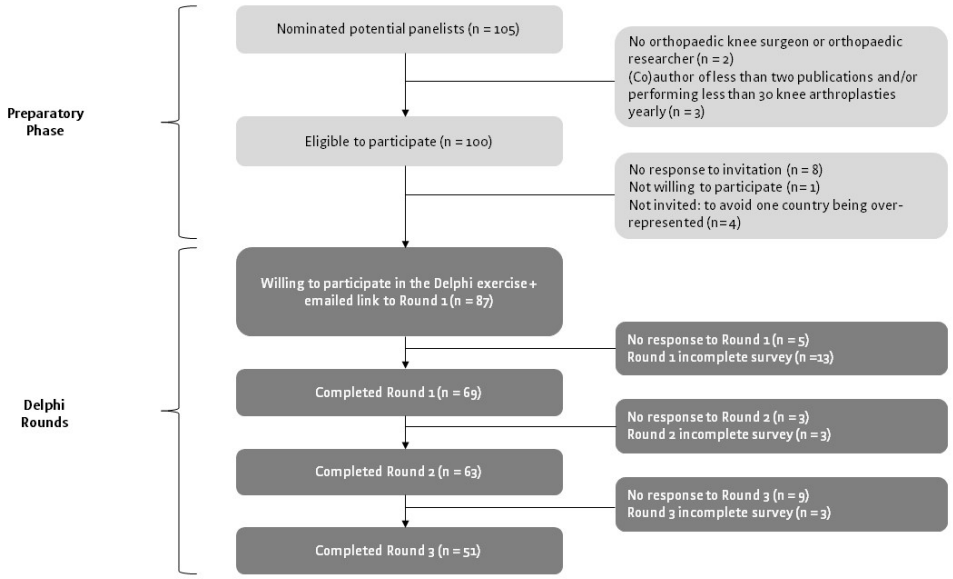


Figure 3. Flowchart of panelists

Round 2

Table 2 (the remaining definitions of the final round) and supplementary table 2 shows the reconsidered scores for face validity and feasibility of the initial definitions and the face validity and feasibility scores for the suggested definitions from round 1. The expert group decided based on consensus that 17 out of a total of 34 definitions with a median score lower than 6.5 for face validity would be removed from the list that served as input for round 3. On the basis of comments of panelists and after discussion among members of the expert advisory group, some adjustments (e.g. 'since before TKA' instead of 'since TKA' or 'no change' (on a scale of: much worse – much better) was defined as poor responder) were made in the wording of 7 definitions. Panelists took on average 17 min time to complete Round 2.

Round 3

The ranking of the 17 remaining definitions in the final round is shown in Table 2. Furthermore, the percentage of panelists that scored at least 1 point for a definition is also displayed. Panelists took on average 9 min time to complete Round 3.

Table 2. Face validity and Feasibility scores and Ranking list of definitions of poor response to TKA after Delphi round 3

Ranking based on total ranking points			Delphi round 2			Delphi round 3	
Ranking	Definition	Underlying domain(s)	Face validity mean (SD)	Feasibility mean (SD)	Sum score Face validity + Feasibility	Total ranking points	N respondents with ≥ 1 point[a] (n = 51)
1	No improvement on transition question on change in pain OR daily knee functioning[b,c,d]	pain, physical functioning	6.5 (1.9)	7.5 (1.5)	14	544	46 (90.2%)
2	Single item question on satisfaction with the outcome (scale: very unsatisfied - very satisfied) Poor responder = very unsatisfied, unsatisfied	satisfaction	7.5 (1.5)	8.5 (1.2)	16	497	47 (92.2%)
3	No improvement on transition question on change in pain[b,c]	pain	7.4 (1.7)	8.1 (1.4)	15.5	386	44 (86.3%)
4	Single item question: "Considering your outcome, are you happy that you had your TKA surgery?" (scale: yes/no) Poor responder = if no	satisfaction	7.1 (2.2)	8.1 (1.6)	15.2	385	43 (84.3%)
5	No improvement on transition question on change in daily knee functioning (rising from sitting, walking, stair climbing)[b,d]	physical functioning	6.6 (1.7)	7.5 (1.7)	14.1	324	43 (84.3%)
6	Single item question on willingness to do TKA surgery again (yes/no) Poor responder = if no	satisfaction	6.9 (2.3)	8.2 (1.7)	15.1	320	40 (78.4%)
7	OKS pain & functioning (scale: 0-48) absolute improvement ≤6	pain, physical functioning	6.8 (1.4)	6.8 (1.6)	13.6	283	34 (66.7%)
8	OMERACT-OARSI responder criteria (WOMAC pain & functioning and global score): Non-responder: (<50% improvement and <20 absolute change in either pain or function) OR (no improvement in at least 2 of the 3 following: <20% improvement and <10 absolute change in either pain, function or patient's global assessment)	pain, physical functioning, global assessment	6.3 (1.8)	4.8 (1.7)	11.1	278	38 (74.5%)
9	Single item question on fulfillment of TKA expectations (scale: to a great extent - not at all) Poor responder = very little, not at all	satisfaction	6.9 (1.6)	7.5 (1.5)	14.4	276	38 (74.5%)
10	OKS PASS <30 (scale: 0-48) PASS: Patient Acceptable Symptom State	pain, physical functioning	6.5 (1.7)	6.6 (1.8)	13.1	257	38 (74.5%)
11	NRS pain >40 in the treated knee (scale: 0-100)	pain	6.8 (1.5)	8.0 (1.5)	14.8	255	40 (78.4%)
12	OKS pain & functioning <26 (scale: 0-48)	pain, physical functioning	6.9 (2.0)	6.7 (1.9)	13.6	244	38 (74.5%)
13	No improvement on transition question on change in knee functioning during moderate activities (gardening, shopping, cycling)[b,d]	physical functioning	6.5 (1.7)	7.4 (1.6)	13.9	235	41 (80.4%)
14	OKS pain & functioning (scale: 0-48) absolute improvement ≤5	pain, physical functioning	7.1 (1.5)	6.9 (1.6)	14	234	36 (70.6%)
15	WOMAC pain, stiffness & functioning (scale: 0-100) absolute improvement <10	pain, knee function, physical functioning	6.1 (1.6)	5.2 (1.7)	11.3	214	37 (72.5%)
16	New KSS symptoms subscale (scale: 0-25) absolute improvement <15	pain	6.2 (1.5)	6.2 (1.7)	12.4	211	34 (66.7%)
17	Single item question on nocturnal knee pain causing sleep disturbance (yes/no) Poor responder = if yes	pain	7.1 (2.3)	8.4 (1.6)	15.5	157	32 (62.7%)

Abbreviations: KSS, Knee Society Score; NRS, Numeric Rating Scale; OKS, Oxford Knee Score; OMERACT-OARSI, Outcome Measures in Arthritis Clinical Trials-Osteoarthritis Research Society International; PASS, Patient Acceptable Symptom State; SD, Standard Deviation; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index

The ranking of definitions is based on total ranking points (column 7);
[a] The number of panelists (%) that voted with at least 1 point for a definition;
[b] Specification of definitions based on transition question(s): How have your pain symptoms or daily knee functioning changed since your TKA?;
[c] Transition questions on change in pain and daily knee functioning range of 1 to 7, with 1 representing very deteriorated and 7 representing very improved. A score <4 was categorized as 'poor response';
[d] Transition question on how daily knee functioning or functioning during moderate activities changed (scale: much worse - much better) Poor responder = much worse, worse, a little worse, unchanged

6

Top Prioritized Definitions

A definition on (dis)satisfaction with the outcome of TKA obtained the highest scores for face validity and feasibility (7.5, and 8.5 respectively), and a definition reflecting change since pre-operative status from the perception of patients (i.e., their opinion on change in pain, daily knee functioning) was highest prioritized (544 points). Furthermore, the sum score of face validity plus feasibility was also shown in table 2. Definitions with >300 total ranking points all had a high sum score (i.e., mean face validity score plus mean feasibility score). Conversely, it was not applicable as some definitions with high sum scores had a low ranking (i.e. definition on ranking position 17).

Discussion

This study is the first to identify and rank potential definitions that may identify poor outcome one year after TKA. The top 6 definitions merits further analysis, as these definitions were both ranked highest and had high sum scores for face validity plus feasibility. Three are transition questions (reflecting change compared to pre-operative status from the perception of patients) and three are single item questions on patient's global appraisal of the outcome. All six definitions are patient-centered, and not clinician-centered. Furthermore, panelists did not favor definitions that rely on existing patient-reported outcome measure (PROM) scores as these definitions received lower scores for feasibility.

The definition on patient (dis)satisfaction with the outcome of TKA scored highest for face validity and feasibility. The highest score for face validity suggests that poor response after TKA is best reflected in this overarching concept. However, knee OA with pain and functional limitations, corroborated by radiographic findings are the indicators for TKA surgery.[20,21] Interestingly, the highest face validity scores for definitions quantifying (changes in) pain and daily functioning were expected. Our results suggest that the concept patient dissatisfaction may capture more than only pain and daily functioning and better reflects "poor response" according to the panelists. Previous studies have demonstrated that persistent pain and functional limitations are the leading reasons for dissatisfaction after TKA, but also that a subset of patients view satisfaction as an evaluation of the process by which care is delivered.[5,6,22] Patient (dis)satisfaction with the process of care delivery may need to be distinguished from (dis)satisfaction with the outcome, as one may not relate directly to the other.[23] There is a widely reported variation in dissatisfaction rates[5], and this variation may in part be explained due to the format of the question[24,25], (e.g. yes/no format, and dichotomized Likert scales or Numeric Rating Scales with variable cut-offs).[5,7] Moreover, in general, single item questions[7] are being used because validated patient (dis)satisfaction questions with standardized response options are scarce.[24] Clement et al. previously demonstrated that the wording of the satisfaction question significantly influences the rate of patient satisfaction one year after TKA.[24] However, despite highest face validity and feasibility, patient (dis) satisfaction was ranked second after a transition question on perceived change in pain and daily knee functioning. The complexity of interpreting patient (dis)satisfaction possibly may explain why panelists preferred a transition question on perceived change in pain and daily knee functioning above a single item question on (dis)satisfaction.

The prioritized list also contains a number of definitions that include a predefined minimal clinical important difference (MCID), which is a metric that represents the minimum difference in the scoring measure that the patient perceives as beneficial or harmful after treatment compared with those who perceived no change.[26] However, definitions containing a MCID received a lower ranking than definitions based on a transition questions. A possible explanation for this is that the change in PROM scores depends on the patient initial baseline status[27], and thus requires preoperative as well as postoperative assessment of PROMs. On the other hand, definitions based on transition questions (including questions on (dis)satisfaction) are subject to recall bias, because patients might not remember their preoperative conditions adequately one year after the procedure.

Prioritized definitions in this study mainly describe change from the patient perspective on underlying domains such as pain, physical functioning, and satisfaction (Table 2). It is noteworthy that the list of ranked definitions do not contain clearly defined, more objective elements as for example knee flexion <90^0, flexion contracture >10^0 or revision surgery within one year after the initial procedure, despite the inclusion of such objective measures in the initial list of definitions. This finding implies that researchers and clinicians place greater emphasis on subjective measures from the patient perspective rather than relying solely on objective measures or the clinical judgment of clinicians.

Perhaps unsurprisingly, the definition on satisfaction with the outcome of TKA received the highest feasibility score. Cost-free availability and brevity makes this definition feasible to measure poor response to TKA. However, it is important to acknowledge that this definition serves as a crude indicator, offering abstract information. While this may be adequate for clinical practice as starting question to elicit problems, it may not provide sufficient detail for research purposes and quality improvement.

Another conclusion that can be drawn from the prioritized list is that the top 6 definitions do not include validated PROMs. Moreover, definitions that do rely on PROMs received lower scores for feasibility. Feasibility considerations motivated by panelists and members of the expert advisory group indicate that an international definition should not depend on previously validated questionnaires as these are not available in all languages and are not easy to assess worldwide and in clinical practice. Furthermore, the volume of questions in PROMs can easily become burdensome. A possible explanation is that validated PROMs are not (yet) feasible for clinical practice or bench marking but more suitable for research purposes.

Strengths & limitations

The strength of a web-based survey is that it ensured anonymity between panelists, which minimizes social pressures and avoids group decisions being dominated by specific experts.[13] Remote data collection facilitated inclusion of a broad range of international key experts in the orthopedic field, with at least 23 different countries being represented.

The main limitation of the present study might be a suboptimal representation of the expert advisory group and Delphi panel, as it did not involve TKA patients or other stakeholders (e.g. allied health practitioners). We deliberately chose not to include patient representatives in this study considering the need for strong English language skills due to the international nature of the study, as well as the complexity associated with the Delphi exercise itself. However, we

processed patient input from the previous interview study and decided to perform a separate study on the prioritization of adverse consequences of TKA among patients.

Another limitation concerns the threshold for the inclusion of definitions in round 3 which was not pre-defined but determined by consensus within the expert advisory group for pragmatic reasons based on the results of round 2. Additionally, several panelists indicated that they were not familiar with certain PROMs or metrics (e.g., MCID, PASS: Patient Acceptable Symptom State) used in the definitions, which could have affected the assessment and ranking of definitions. After round 2 some newly suggested definitions have been reformulated as several panelist commented that some definitions were not completely clear, which could also have affected the assessment of face validity and feasibility by panelists. Reasons for non-response of the different rounds were not provided.

Despite our comprehensive efforts to recruit panelists from around the world, there was under-representation of several continents. Most of the panelists worked in a European country, North America or Australia, which may limit generalizability of the findings. The main contributing factor to this is that the Delphi panel was set up by the members of the expert advisory group working on these continents.

Conclusions

This study with representation from 23 countries across the globe is the first to attempt to define poor response to TKA. It is essential to be able to identify these patients, for future quality improvement efforts. We identified seventeen potential definitions, of which six definitions we believe justify further study and potential implementation in quality assessment studies. All six definitions are patient-centered. Three are transition questions (reflecting change compared to pre-operative status from the perception of patients) and three are single item questions on patient satisfaction with the outcome. Remarkably, none are based on assessment of knee function by the clinician and none are complication, surgery- or revision-related. These six definitions for measuring poor response to TKA merit further analysis.

Acknowledgements

The authors thank Andrew J. Porteous for his participation in de expert advisory group and contribution to the study. We would like to thank the following panelists for the time and effort to participate in the Delphi study: Maren Falch Lindberg, A/Prof & Senior Researcher, University of Oslo & Lovisenberg Diaconal Hospital; Ponky Firer, Prof & Senior Consultant, Linksfield Knee Clinic; Michael Whitehouse, Prof, University of Bristol; Jore Willems, Orthopaedic Surgeon, Sint Maartenskliniek; Joris Jansen, Orthopaedic Surgeon, Alrijne Hospital; Bernardo Innocenti, Prof, Université Libre de Bruxelles, Andrew Price, Prof & Clinical Director & Knee Surgeon, Oxford University, Nuffield Orthopaedic Centre; Vikki Wylde, Prof, University of Bristol; Luca Matascioli, Knee surgeon, Fondazione Poliambulanza; Stergios Lazarinis, Head of Dep, Uppsala University Hospital; Frank-Christiaan Wagenaar, Orthopaedic consultant, OCON Orthopedic Clinic; Ran Schwarzkopf, Prof, NYU Hospital for Joint Diseases, NYU Langone Orthopedic Hospital; Ashok Rajgopal, Group Chairman, Medanta Institute of Musculoskeletal Disorders; Thomas Paszicsnyek, Orthopaedic Surgeon, Medfit; Jon Goosen, Orthopaedic Surgeon, Sint Maartenskliniek; Lex Boerboom, Orthopaedic Surgeon, University Medical Center Groningen; Johannes Beckmann, Prof and Head of Clinic, Hospital Barmherzige Brüder München; Derk van Kampen, Orthopaedic Surgeon, Dijklander Hospital; Stefano Marco Paolo Rossi, Knee surgeon, Fondazione Poliambulanza; Sebastiaan van der Groes, Orthopaedic surgeon, Radboud University Medical Center; Michael Dunbar, Prof, Dalhousie University; Federica Rosso, Orthopaedic Surgeon, AO Ordine Mauriziano Hospital; David Hamilton, Lecturer/Principle Investigator, Edinburgh Napier University; Søren Skou, Prof, University of Soutern Denmark and Næstved, Slagelse and Ringsted Hospital, Rob Janssen, Prof & A/Prof & Orthopaedic Surgeon, Máxima Medical Center & University of Technology & Fontys University of Applied Sciences; Arild Aamodt, Orthopaedic Surgeon, Lovisenberg Diaconal Hospital; Stefan de Boer, Orthopaedic Surgeon, VieCuri Medical Center; Rob Nelissen, Chair & Prof, Leiden University Medical Center; Nicolaas Budhiparafma, Prof & A/Prof, Universitas Airlangga, Leiden University Medical Center; Jaap Tolk, Orthopaedic Surgeon, Erasmus MC Sophia Children's Hospital; Jean-Noel Argenson, Chair & Prof, The institute for Locomotion; Corné van Loon, Orthopaedic Surgeon, Rijnstate Hospital; Andrea Baldini, Institute Director, IFCA Clinic; Carsten Tibesku, Prof & Orthopaedic Surgeon, KniePraxis; Bryan Springer, Prof & Fellowship Director, OrthoCarolina Hip and Knee Center; Justine Naylor, Conjoint A/Prof & Senior Principal Research Fellow, SWSLHD; Ilana Ackerman, Prof, Monash University; Bas Fransen, Arthroplasty Fellow, University of British Columbia; Siegfried Hofmann, Prof & Head, Knee Training Center Stolzalpe; Jan de Vos, Orthopaedic Surgeon, Wilgers Hospital; Thomas Heyse, Prof & Orthopaedic Surgeon, Red Cross Hospital; Henrik Schrøder, A/prof & Head of Research & Senior Consultant, Naestved Hospital; Bernhard Christen, Orthopaedic Surgeon, articon; Thomas Luyckx, Knee Surgeon & Visiting Prof, AZ Delta & KULeuven; Andrew Toms, Orthopaedic Surgeon & Clinical Director & Academic Head, Princess Elizabeth Orthopaedic Centre & Royal Devon & Exeter Hospital; Karin de Kroon, Orthopaedic Surgeon, Gelre Hospital Apeldoorn; Hans-Peter van Jonbergen, Orthopaedic Surgeon, Deventer Hospital; Peter Feczko, Orthopaedic Surgeon, Maastricht University Medical Center; Fabian Poletti, Orthopaedic Surgeon, Nykøbing Falster Hospital; Pieter van Driel, Orthopaedic Surgeon, Isala Hospital Zwolle; Lucien Keijser, Orthopaedic Surgeon, Northwest Clinics Alkmaar. Note that not all panelists gave consent to be included in the acknowledgements.

References

1. Pabinger C, Lothaller H, Geissler A. Utilization rates of knee-arthroplasty in OECD countries. Osteoarthritis Cartilage 2015;23(10):1664-73.
2. OECD. Health at a Glance. 2019. OECD Indicators.
3. Klug A, Gramlich Y, Rudert M, et al. The projected volume of primary and revision total knee arthroplasty will place an immense burden on future health care systems over the next 30 years. Knee Surg Sports Traumatol Arthrosc 2021;29(10):3287-3298.
4. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010;468(1):57-63.
5. DeFrance M, Scuderi G. Are 20% of Patients Actually Dissatisfied Following Total Knee Arthroplasty? A Systematic Review of the Literature. J Arthroplasty 2022;
6. Halawi MJ, Jongbloed W, Baron S, Savoy L, Williams VJ, Cote MP. Patient Dissatisfaction After Primary Total Joint Arthroplasty: The Patient Perspective. J Arthroplasty 2019;34(6):1093-1096.
7. Te Molder MEM, Smolders JMH, Heesterbeek PJC, van den Ende CHM. Definitions of poor outcome after total knee arthroplasty: an inventory review. BMC Musculoskelet Disord 2020;21(1):378.
8. Price AJ, Longino D, Rees J, et al. Are pain and function better measures of outcome than revision rates after TKR in the younger patient? Knee 2010;17(3):196-9.
9. Singh JA, Dowsey MM, Dohm M, et al. Achieving Consensus on Total Joint Replacement Trial Outcome Reporting Using the OMERACT Filter: Endorsement of the Final Core Domain Set for Total Hip and Total Knee Replacement Trials for Endstage Arthritis. J Rheumatol 2017;44(11):1723-1726.
10. Goodman SM, Mehta BY, Mandl LA, et al. Validation of the Hip Disability and Osteoarthritis Outcome Score and Knee Injury and Osteoarthritis Outcome Score Pain and Function Subscales for Use in Total Hip Replacement and Total Knee Replacement Clinical Trials. J Arthroplasty 2020;35(5):1200-1207.e4.
11. Goodman SM, Mehta BY, Kahlenberg CA, et al. Assessment of a Satisfaction Measure for Use After Primary Total Joint Arthroplasty. J Arthroplasty 2020;35(7):1792-1799.e4.
12. Singh JA, Mehta B, Mirza SZ, et al. When Has a Knee or Hip Replacement Failed? A Patient Perspective. J Rheumatol 2021;48(3):447-453.
13. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. Palliat Med 2017;31(8):684-706.
14. Veugelers R, Gaakeer MI, Patka P, Huijsman R. Improving design choices in Delphi studies in medicine: the case of an exemplary physician multi-round panel study with 100% response. BMC Med Res Methodol 2020;20(1):156.
15. Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol 2014;67(4):401-9.
16. SurveyMonkey. 2021. Retrieved from: https://www.surveymonkey.com. Date last accessed: 12 April 2021.
17. Murphy MK, Black NA, Lamping DL, et al. Consensus development methods, and their use in clinical guideline development. Health Technology Assessment 1998;2(3):1-88.
18. Sadler GR, Lee HC, Lim RS, Fullerton J. Recruiting hard-to-reach United States population sub-groups via adaptations of the snowball sampling strategy. Nurs Health Sci 2010;12(3):369-74.
19. Trevelyan EG, Robinson N. Delphi methodology in health research: how to do it? European Journal of Integrative Medicine 2015;7(4):423-428.
20. Cross WW, Saleh KJ, Wilt TJ, Kane RL. Agreement about indications for total knee arthroplasty. Clin Orthop Relat Res 2006;446:34-9.

21. Gademan MG, Hofstede SN, Vliet Vlieland TP, Nelissen RG, Marang-van de Mheen PJ. Indication criteria for total hip or knee arthroplasty in osteoarthritis: a state-of-the-science overview. BMC Musculoskelet Disord 2016;17(1):463.
22. Rodriguez-Merchan EC. Patient Satisfaction Following Primary Total Knee Arthroplasty: Contributing Factors. Arch Bone Jt Surg 2021;9(4):379-386.
23. Donabedian A. The quality of care. How can it be assessed? Jama 1988;260(12):1743-8.
24. Clement ND, Bardgett M, Weir D, Holland J, Gerrand C, Deehan DJ. The rate and predictors of patient satisfaction after total knee arthroplasty are influenced by the focus of the question: a standard satisfaction question is required. Bone Joint J 2018;100-b(6):740-748.
25. Khatib Y, Badge H, Xuan W, Naylor JM, Harris IA. Patient satisfaction and perception of success after total knee arthroplasty are more strongly associated with patient factors and complications than surgical or anaesthetic factors. Knee Surg Sports Traumatol Arthrosc 2020;28(10):3156-3163.
26. Clement ND, Bardgett M, Weir D, Holland J, Gerrand C, Deehan DJ. What is the Minimum Clinically Important Difference for the WOMAC Index After TKA? Clin Orthop Relat Res 2018;476(10):2005-2014.
27. Copay AG, Subach BR, Glassman SD, Polly DW, Jr., Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J 2007;7(5):541-6.

6

Supplementary material

Supplementary table 1. Initial draft list of definitions (Round 1)

Definition
NRS pain >40 (scale: 0-100)
IKSS pain <30 (scale: 0-50)
Knee flexion <90°
IKSS functioning <60 (scale: 0-100)
Single item question: "Have you been able to return to the activity (or activities) that your knee stopped you from doing one year ago?" (scale: yes/no) Poor responder = if no
Single item question on satisfaction with the outcome (scale: very unsatisfied - very satisfied) Poor responder = very unsatisfied, unsatisfied
OKS pain & functioning (scale: 0-48) absolute improvement ≤5
OKS pain & functioning (scale: 0-48) absolute improvement ≤6
OMERACT-OARSI responder criteria (WOMAC pain & functioning and global score): Non-responder: (<50% improvement and <20 absolute change in either pain or function) OR (no improvement in at least 2 of the 3 following: <20% improvement and <10 absolute change in either pain, function or patient's global assessment)
WOMAC pain, stiffness & functioning (scale: 0-100) absolute improvement <7.5
< 50% improvement and an absolute change of < 20 in pain scale 0-100 (yes/no) Poor responder = if yes
No improvement on transition question on change in pain[1]
< 30% improvement and an absolute change of < 10 in knee functioning scale 0-100 (yes/no) Poor responder = if yes
< 30% improvement and an absolute change of < 10 in knee functioning scale 0-100 (yes/no) Poor responder = if yes
No improvement on transition question on change in daily knee functioning (rising from sitting, walking, stair climbing)[2]
No improvement on transition question on change in knee functioning during moderate activities (gardening, shopping, cycling)[2]
Composite question: "Have you had any serious complication (e.g. infection) that has required further surgery or revision of the prosthesis?" (scale: yes/no) OR "Has your knee been replaced?" (scale: yes/no) Poor responder = if yes on one or both questions
< 50% improvement and an absolute change of < 20 in pain scale 0-100 (yes/no) OR < 30% improvement and an absolute change of < 10 in knee functioning scale 0-100 (yes/no) Poor responder = if yes on one or both questions
No improvement on transition question on change in pain OR daily knee functioning[1,2]
< 50% improvement and an absolute change of < 20 in pain scale 0-100 (yes/no) OR No improvement on transition question on change in pain OR daily knee functioning[1,2] Poor responder = if yes on the first question or much worse, worse, a little worse on the second question
< 30% improvement and an absolute change of < 10 in knee functioning scale 0-100 (yes/no) OR No improvement on transition question on change in pain OR daily knee functioning[1,2] Poor responder = if yes on the first question or much worse, worse, a little worse on the second question
< 20% improvement and an absolute change of < 10 in pain (yes/no) OR < 20% improvement and an absolute change of < 10 in knee functioning (yes/no) AND No improvement on transition question on change in pain OR daily knee functioning[1,2] Poor responder = if yes on one of the first two questions and much worse, worse, a little worse on the third question
< 50% improvement and an absolute change of < 20 in pain (yes/no) OR How happy/satisfied are you with the level of improvement in your pain? Poor responder = if yes on the first question or very unsatisfied, unsatisfied on the second question
< 50% improvement and an absolute change of < 20 in pain (yes/no) OR How happy/satisfied are you with the level of improvement in your pain? OR How happy/satisfied are you with the level of improvement in your knee functioning? Poor responder = if yes on the first question or much worse, worse, a little worse on the second question or very unsatisfied, unsatisfied on the third question
< 30% improvement and an absolute change of < 10 in knee functioning (yes/no) OR How happy/satisfied are you with the level of improvement in your knee functioning? Poor responder = if yes on the first question or very unsatisfied, unsatisfied on the second question
< 30% improvement and an absolute change of < 10 in knee functioning (yes/no) OR How happy/satisfied are you with the level of improvement in your pain? OR How happy/satisfied are you with the level of improvement in your knee functioning? Poor responder = if yes on the first question or very unsatisfied, unsatisfied on the second and third question

Abbreviations: KSS, Knee Society Score; NRS, Numeric Rating Scale; OKS, Oxford Knee Score; OMERACT-OARSI, Outcome Measures in Arthritis Clinical Trials-Osteoarthritis Research Society International; PASS, Patient Acceptable Symptom State; SD, Standard Deviation; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index

[1] Transition questions on change in pain and daily knee functioning range of 1 to 7, with 1 representing very deteriorated and 7 representing very improved. A score <4 was categorized as 'poor response';
[2] Transition question on how daily knee functioning or functioning during moderate activities changed (scale: much worse - much better) Poor responder = much worse, worse, a little worse, unchanged

6

Supplementary table 2. Complete overview of face validity and feasibility scores of Delphi round 1 and 2

Definitions Delphi Round 1 and 2	Round 1		Round 2	
	Face validity mean (SD)	Feasibility mean (SD)	Face validity mean (SD)	Feasibility mean (SD)
NRS pain >40 (scale: 0-100)	6.5 (1.9)	8.1 (2.0)	6.8 (1.5)	8.0 (1.5)
IKSS pain <30 (scale: 0-50)	6.4 (2.1)	7.0 (2.1)	6.2 (1.5)	6.2 (1.7)
Knee flexion <90°	5.9 (2.4)	7.5 (2.4)	5.9 (2.1)	7.3 (1.7)
IKSS functioning <60 (scale: 0-100)	6.2 (2.0)	6.9 (2.0)	5.8 (1.4)	6.1 (1.6)
Single item question: "Have you been able to return to the activity (or activities) that your knee stopped you from doing one year ago?" (scale: yes/no) Poor responder = if no	5.4 (2.5)	7.2 (2.3)	5.1 (1.6)	6.7 (1.7)
Single item question on satisfaction with the outcome (scale: very unsatisfied - very satisfied) Poor responder = very unsatisfied, unsatisfied	7.4 (2.3)	8.4 (1.8)	7.5 (1.5)	8.5 (1.2)
OKS pain & functioning (scale: 0-48) absolute improvement ≤5	7.1 (1.8)	7.6 (1.9)	7.1 (1.5)	6.9 (1.6)
OKS pain & functioning (scale: 0-48) absolute improvement ≤6	6.7 (1.9)	7.4 (2.0)	6.8 (1.4)	6.8 (1.6)
OMERACT-OARSI responder criteria (WOMAC pain & functioning and global score): Non-responder: (<50% improvement and <20 absolute change in either pain or function) OR (no improvement in at least 2 of the 3 following: <20% improvement and <10 absolute change in either pain, function or patient's global assessment)	6.6 (2.4)	5.5 (2.5)	6.3 (1.8)	4.8 (1.7)
WOMAC pain, stiffness & functioning (scale: 0-100) absolute improvement <7.5	6.2 (2.3)	6.6 (2.3)	6.1 (1.6)	5.2 (1.7)
< 50% improvement and an absolute change of < 20 in pain scale 0-100 (yes/no) Poor responder = if yes	6.1 (2.0)	6.7 (2.2)	5.7 (1.7)	6.4 (1.9)
No improvement on transition question on change in pain[1]	7.1 (2.3)	8.2 (2.0)	7.4 (1.7)	8.1 (1.4)
< 30% improvement and an absolute change of < 10 in knee functioning scale 0-100 (yes/no) Poor responder = if yes	5.2 (2.3)	6.1 (2.5)	5.3 (1.7)	5.8 (1.6)
No improvement on transition question on change in daily knee functioning (rising from sitting, walking, stair climbing)[2]	6.7 (2.1)	7.8 (1.9)	6.6 (1.7)	7.5 (1.7)
No improvement on transition question on change in knee functioning during moderate activities (gardening, shopping, cycling)[2]	6.5 (2.2)	7.8 (2.0)	6.5 (1.7)	7.4 (1.6)
Composite question: "Have you had any serious complication (e.g. infection) that has required further surgery or revision of the prosthesis?" (scale: yes/no) OR "Has your knee been replaced?" (scale: yes/no) Poor responder = if yes on one or both questions	5.9 (3.2)	7.6 (2.6)	5.2 (2.4)	7.7 (1.8)
< 50% improvement and an absolute change of < 20 in pain scale 0-100 (yes/no) OR < 30 % improvement and an absolute change of < 10 in knee functioning scale 0-100 (yes/no) Poor responder = if yes on one or both questions	5.5 (2.2)	5.8 (2.3)	5.4 (1.6)	5.4 (1.8)
No improvement on transition question on change in pain OR daily knee functioning[1]	6.8 (2.3)	7.6 (2.2)	6.5 (1.9)	7.5 (1.5)
< 50% improvement and an absolute change of < 20 in pain scale 0-100 (yes/no) OR No improvement on transition question on change in pain OR daily knee functioning[1,2] Poor responder = if yes on the first question or much worse, worse, a little worse on the second question	5.7 (2.2)	6.7 (2.0)	5.8 (1.7)	6.2 (1.5)
< 30 % improvement and an absolute change of < 10 in knee functioning scale 0-100 (yes/no) OR No improvement on transition question on change in pain OR daily knee functioning[1,2] Poor responder = if yes on the first question or much worse, worse, a little worse on the second question	5.6 (2.1)	6.5 (2.1)	5.4 (1.7)	6.1 (1.5)
< 20% improvement and an absolute change of < 10 in pain (yes/no) OR < 20% improvement and an absolute change of < 10 in knee functioning (yes/no) AND No improvement on transition question on change in pain OR daily knee functioning[1,2] Poor responder = if yes on one of the first two questions and much worse, worse, a little worse on the third question	5.6 (2.5)	6.2 (2.6)	5.4 (2.1)	5.6 (1.9)
< 50% improvement and an absolute change of < 20 in pain (yes/no) OR How happy/satisfied are you with the level of improvement in your pain? Poor responder = if yes on the first question or very unsatisfied, unsatisfied on the second question	5.8 (2.2)	6.7 (2.4)	5.9 (1.9)	6.2 (1.8)
< 50% improvement and an absolute change of < 20 in pain (yes/no) OR How happy/satisfied are you with the level of improvement in your pain? OR How happy/satisfied are you with the level of improvement in your knee functioning? Poor responder = if yes on the first question or much worse, worse, a little worse on the second question or very unsatisfied, unsatisfied on the third question	5.7 (2.3)	6.3 (2.5)	5.4 (2.0)	5.5 (1.9)
< 30% improvement and an absolute change of < 10 in knee functioning (yes/no) OR How happy/satisfied are you with the level of improvement in your knee functioning? Poor responder = if yes on the first question or very unsatisfied, unsatisfied on the second question	5.6 (2.2)	6.5 (2.3)	5.3 (1.9)	5.7 (1.8)

< 30% improvement and an absolute change of < 10 in knee functioning (yes/no) OR How happy/satisfied are you with the level of improvement in your pain? OR How happy/satisfied are you with the level of improvement in your knee functioning? Poor responder = if yes on the first question or very unsatisfied, unsatisfied on the second and third question	5.7 (2.3)	6.4 (2.2)	5.2 (2.0)	5.5 (1.9)
Single item question: "Considering your outcome, are you happy that you had your TKA surgery?" (scale: yes/no) Poor responder = if no			7.1 (2.2)	8.1 (1.6)
Lack of extension > 10° (scale: yes/no) Poor responder = if yes			5.2 (2.5)	6.6 (2.3)
Single item question on willingness to do TKA surgery again (scale: yes/no) Poor responder = if no			6.9 (2.3)	8.2 (1.7)
Single item question on fulfillment of TKA expectations (scale: to a great extent - not at all) Poor responder = very little, not at all			6.9 (1.6)	7.5 (1.5)
OKS PASS <30 (scale: 0-48) PASS: Patient Acceptable Symptom State			6.5 (1.7)	6.6 (1.8)
OKS pain & functioning <26 (scale: 0-48)			6.9 (2.0)	6.7 (1.9)
Composite question: "Are you dissatisfied with either pain OR function? (scale: yes/no) Poor responder = if yes for one or both questions			6.2 (2.1)	6.8 (2.1)
Single item question on nocturnal knee pain causing sleep disturbance (scale: yes/no) Poor responder = if yes			7.1 (2.3)	8.4 (1.6)
Single item question: "Are you aware of your knee every day?" (yes/no) Poor responder = if yes			4.9 (2.1)	7.0 (2.1)

Abbreviations: KSS, Knee Society Score; NRS, Numeric Rating Scale; OKS, Oxford Knee Score; OMERACT-OARSI, Outcome Measures in Arthritis Clinical Trials-Osteoarthritis Research Society International; PASS, Patient Acceptable Symptom State; SD, Standard Deviation; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index

[1] Transition questions on change in pain and daily knee functioning range of 1 to 7, with 1 representing very deteriorated and 7 representing very improved. A score <4 was categorized as 'poor response';
[2] Transition question on how daily knee functioning or functioning during moderate activities changed, (scale: much worse - much better) Poor responder = much worse, worse, a little worse, unchanged

Supplementary table 3. Round 2 new definitions. New definitions generated from panelists' Round 1 free-text responses and included in the Round 2 survey

Definition
Single item question: "Considering your outcome, are you happy that you had your TKA surgery?" (scale: yes/no) Poor responder = if no
Lack of extension > 10° (scale: yes/no) Poor responder = if yes
Single item question on willingness to do TKA surgery again (scale: yes/no) Poor responder = if no
Single item question on fulfillment of TKA expectations (scale: to a great extent - not at all) Poor responder = very little, not at all
OKS PASS <30 (scale: 0-48) PASS: Patient Acceptable Symptom State
OKS pain & functioning <26 (scale: 0-48)
Composite question: "Are you dissatisfied with either pain OR function? (scale: yes/no) Poor responder = if yes for one or both questions
Single item question on nocturnal knee pain causing sleep disturbance (scale: yes/no) Poor responder = if yes
Single item question: "Are you aware of your knee every day?" (yes/no) Poor responder = if yes

Chapter 7

Prioritization of adverse consequences after total knee arthroplasty contributing to a poor response: A Best-Worst Scaling Exercise among total knee arthroplasty patients and knee specialists

Malou E.M. te Molder
Lise M. Verhoef
José M.H. Smolders
Petra J.C. Heesterbeek
Cornelia H.M. van den Ende

Revised version of this chapter is published as The Journal of Arthroplasty. 2024;39(3): 651-7

Abstract

Background

Total knee arthroplasty (TKA) can have a number of adverse consequences for patients that might contribute to a poor outcome. This study aimed to prioritize these consequences, from the perspective of patients and knee specialists.

Methods

There were 95 TKA patients and 63 knee specialists who prioritized a set of 29 adverse consequences, based on a previous qualitative study, using a Maximum Difference Scaling method. A hierarchical Bayesian analysis was used to calculate relative importance scores. Differences and agreements between patients versus knee specialists and satisfied versus dissatisfied patients were analyzed using Mann-Whitney-U tests and Kendall's coefficients of concordance.

Results

There were four out of five items in the top-5 of both patients and knee specialists that were similar, however, the ranking was different. The highest-ranked consequence for patients was: "Inability to do normal activities such as walking, cycling, swimming and heavy household chores", while knee specialists ranked: "No improvement in pain during the day" as the highest. "No improvement in walking" was in the patients' top-5, but was not ranked in the top-5 of knee specialists. For satisfied and dissatisfied patients, the top-5 of consequences was similar.

Conclusion

Comparable perspectives were found for patients versus knee specialists and satisfied versus dissatisfied patients on the importance of adverse consequences after TKA. However, when looking in more detail, differences in ranking of specific subitems suggest that patients place slightly more importance on the inability to perform valued activities, while knee specialists prioritize lack of pain relief to a higher degree.

Keywords

Total knee arthroplasty, adverse consequences, patient perspective, outcome, specialist perspective

Introduction

Total knee arthroplasty (TKA) has been widely-established as a successful procedure for advanced symptomatic osteoarthritis (OA). Despite overall clinical improvement[1], 10 to 20% of patients report that they are dissatisfied due to insufficient pain relief and limitations in physical functioning.[2-4] To improve these dissatisfaction rates, a clear and valid definition of what poor outcomes after TKA entails is needed. Currently, various dichotomous definitions comprising one or more outcome dimensions are used to quantify the proportion of patients who have a poor outcome after TKA.[5] This large variety of definitions impedes the comparisons of poor response to TKA over time and across hospitals and countries. While mostly only one outcome dimension is used as primary endpoint, it is recognized that it is necessary to use a combination of outcome domains to accurately capture failure (ie, poor response) after TKA.[6]

Little research has been done on outcome domains that should be included in a definition of poor response. One study focusing on defining total joint arthroplasty (TJA) failure from the patient's perspective used nominal group technique and prioritized refractory pain after TJA as more important than surgical failure (ie, complications, revisions).[7] However, nominated group responses/themes about TJA failure were ranked. The latter study did not provide insight in differences between the ranking of patients and knee specialists about adverse consequences of TKA for patients. A recent qualitative study identified adverse consequences after TKA as perceived by patients and knee specialists.[8] This study showed that knee specialists put emphasis on surgical failure, unexplained pain, and impairments that limit patients' physical functioning, while patients were focused on the process of adapting to the knee, awareness of the artificial knee, and limitations they experience during valued daily activities. Therefore, there are probably differences in what patients and knee specialists consider the most important contributors to poor outcome after TKA. However, adverse consequences of TKA for patients were not previously ranked by knee specialists and compared to the ranking of patients.

It is essential to evaluate whether the consequences of TKA for patients are ranked differently by knee specialists in order to determine if they are aware of the most and least important contributors to poor outcome after TKA according to the patients' perspective. Therefore, the objective of this study was to prioritize postoperative consequences of TKA for patients that contribute to a poor response, according to the perspective of patients and knee specialists.

Methods

Study Design

In this study, a Best-Worst scaling (BWS) questionnaire was used to determine the relative importance of adverse postoperative consequences for TKA patients from a previous qualitative study.[8] Best-worst scaling presents respondents with a series of item subsets from a master set.[9] Respondents were asked to indicate for each subset their most and least important item, and complete a number of these subsets where each subset contained a different random selection of items. From these selections, a respondent's item ranking was constructed.[9] The TKA patients and knee specialists were asked to prioritize numerous consequences after TKA that contribute to poor response. The study (protocol file number

2022-15740) was presented to the Medical Ethical Committee of the region Arnhem-Nijmegen. An exemption was obtained, as ethical approval for this type of study was not required under Dutch law. This study complied with the declaration of Helsinki and all respondents provided digital informed consent.

Participants

No guidance is provided in the literature regarding the minimal sample size for a desired statistical power for BWS methods.[10] Sample sizes of previous studies evaluated in a review ranged between 15 and 9,289 participants.[11] For this study, we aimed to include 100 patients and 50 knee specialists to be able to calculate reliable ranking scores and perform subgroup analyses, while keeping recruitment feasible.

Recruitment of patients

Patients were recruited from the orthopaedic department of Sint Maartenskliniek, Nijmegen, the largest specialized orthopaedic hospital in the Netherlands. Patients who received a TKA 1 to 3 years prior for OA were selected from the electronic patient files, screened for eligibility, and approached for the study with an information letter by e-mail. One reminder was sent in case patients did not respond. An a priori decision was made to limit follow-up from 1 to 3 years, because pain and physical functioning outcome from 1 to 3 years following TKA have been shown to be reasonably stable.[12-14] Furthermore, patients had to be ≥ 18 years, able to communicate well in Dutch, possess basic computer skills and an email address, be consented to be contacted by e-mail, and be willing to participate in the study, as well as sign an (digital) informed consent.

Out of 188 patients that were approached, 133 patients provided informed consent and started the online questionnaire, of whom 102 completed the BWS exercise (Figure 1). There were 95 patients included in the analyses because 7 patients gave inconsistent answers based on a Root Likelihood (RLH) <0.269 and were excluded. Characteristics of 95 patients (response rate 51%) are shown in Table 1.

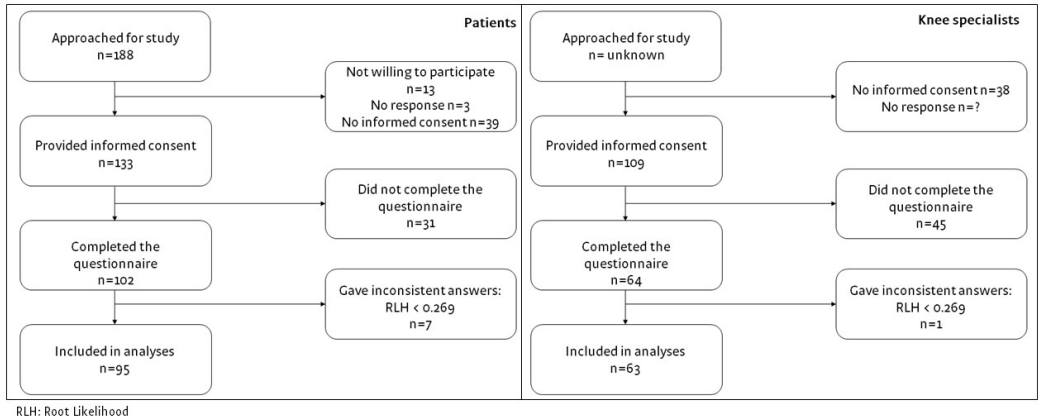


Figure 1. Flowcharts of patients and knee specialists

Table 1. Demographic characteristics of patients and knee specialists

Characteristics	Values
Patients (n = 95)	
Mean age (range)	68 (42 to 89)
Women, n (%)	64 (67)
Level of education[a], n (%)	
No education	1 (1)
Primary	2 (2)
Secondary	53 (56)
Tertiary	39 (41)
Reason for total knee arthroplasty (TKA) according to patients, n (%)	
Osteoarthritis	11 (12)
Pain	18 (19)
Limited knee function	8 (8)
Pain and limited knee function	17 (18)
Pain during walking	33 (35)
Pain during walking and pain at night	7 (7)
Stiff knee	1 (1)
Side of TKA, n (%)	
One-sided, left	17 (18)
One-sided, right	35 (37)
Two-sided	43 (45)
Time since TKA surgery in years, n (%)	
± 1	25 (26)
± 1.5	20 (21)
± 2	24 (25)
> 2	26 (27)
Currently employed, n (%)	25 (26)
Satisfied with TKA, n (%)	
Yes	79 (83)
No	13 (14)
Don't know	3 (3)

7

Knee specialists (n = 63)	
Mean age (range)	45 (25 to 64)
Women, n (%)	27 (43)
Occupation, n (%)	
Orthopaedic surgeon	21 (33)
Physician assistant	10 (16)
Nurse practitioner	9 (14)
Physiotherapist	23 (37)
Setting, n (%)	
General hospital	10 (16)
Top-clinical teaching hospital	16 (25)
Specialized orthopaedic hospital	10 (16)
Academic hospital	5 (8)
Physiotherapy practice	22 (35)
Experience in TKA surgery, treatment and/or rehabilitation in years, median (25^{th}–75^{th} percentile)	12 (8-21)

[a]Primary: primary education; Secondary: lower secondary education, upper secondary education; Tertiary: short-cycle tertiary education, bachelor's or equivalent, master's or equivalent

Recruitment of knee specialists

The following health care professionals (HCPs) with expertise on knee arthroplasty surgery and its rehabilitation were recruited: orthopaedic surgeons specialized in TKA surgery, physician assistants (PA) and nurse practitioners (NP) specialized in orthopaedics, and physiotherapists specialized in TKA rehabilitation. All knee specialists were invited by e-mail; orthopaedic surgeons were invited through the Dutch Knee Society (DKS), PAs and NPs through the working group PA-NP from the Dutch Orthopaedic Association, and physiotherapists through a national OA network for health professionals.

A total of 109 knee specialists provided informed consent and started the questionnaire, of whom 63 finished the BWS exercise. It was not possible to calculate the response rate for knee specialists as they were invited through societies or a national network, therefore it is unknown how many knee specialists actually were approached.

Designing the survey

We developed an online survey consisting of 2 parts. The first part assessed basic demographic characteristics of all respondents: age and sex. Patients were asked additional questions on the knee prosthesis: reason for TKA according to the patients, side, time since TKA surgery, and whether they were satisfied with their TKA. In the survey for knee specialists, we assessed the occupation, the working environment and working experience in TKA surgery and/or care. In the second part of the survey, patients and knee specialists prioritized consequences after TKA according to relative importance (RI) by means of a BWS exercise. Consequences after TKA according to both patients and HCPs from a previous qualitative study by Molder et al were used to construct an initial list of 30 items (consequences). An orthopaedic surgeon, two orthopaedic researchers, a physiotherapist/researcher, a researcher with knowledge in performing BWS studies, and three patient research partners contributed to discuss the wording of the items and refined the initial list to a final list of 29 items that served as input for the BWS exercise. The survey was then pilot-tested in one post-operative patient and one orthopaedic surgeon and evaluated with a researcher (M.t.M.). During the evaluation, the following topics were discussed: clarity of information; response time to complete the survey, and understanding of the list of items. Findings from the evaluations were used to further optimize the survey and item list. Sawtooth Software's Lighthouse Studio (version 9.14.2, Provo, UT) was used to develop the online questionnaire with the BWS exercise. The BWS exercise consisted of 20 unique choice tasks in which five items were shown to ensure that every respondent would rank every item at least 3 times.[9] An example of a BWS question is shown in Figure 2. The software created the optimal design of subsets based on 1,000 iterations. A total of 300 versions were created to ensure a variety of combinations of items and a random order among respondents, to avoid higher importance being given to the first mentioned items.

Analyses

Descriptive analyses were used for demographic characteristics of respondents. The choices made by respondents in the BWS exercise were analyzed using hierarchical Bayesian (HB) estimations to estimate the relative importance (RI) scores. The hierarchical Bayesian estimations allowed us to estimate the individual level of importance by combining information from individuals' specific choices with the distribution of importance across respondents, computing individual-level weights under the logit rule. Raw scores were generated by iterations on an interval scale. To facilitate interpretation, the scores were rescaled to a standardized 0 to 100 ratio scale; the higher the score, the heavier the item weighs. All RIs sum to 100 for each individual. Thus, the RIs represent the relative importance of an item in relation to all other items. Participants that gave inconsistent answers on the BWS questionnaire (Root Likelihood (RLH) below the recommended cut-off of 0.269[15]) were excluded from the analyses. Incomplete surveys were not analyzed, and HB analyses were performed for patients and knee specialists separately. The three undecisive patients regarding (dis)satisfaction with their TKA were included in the dissatisfied patient group for the analyses.

Mann-Whitney-U tests were used to analyze differences in RIs between patients and knee specialists. Possible associations in the ranking of RIs between patients versus knee specialists and between satisfied patients versus dissatisfied patients were explored by using Kendall's coefficient of concordance (Kendall's W). Analyses were performed with Stata 17 software and SPSS statistics version 26.0.0 (IBM Corporation, Armonk, NY).

Display of data

Items with the largest differences in scores between patients versus knee specialists and satisfied versus dissatisfied patients were displayed in a table.

Please read the five items below. Indicate which item is the **most important** contributor, and which is the **least important** contributor to poor outcome after TKA.

12 / 20

Most important		Least important
○	Continued awareness of the knee prosthesis	○
○	No improvement in pain with weight bearing/during activities	○
○	No improvement in pain at night	○
○	No improvement in climbing stairs	○
○	Inability to do normal activities such as walking, cycling, swimming, heavy household chores (window washing and gardening)	○

Back Next

0% 100%

Figure 2. Example of a BWS question

Results

Most important adverse consequences in TKA patients

Figure 3 shows the top-5 highest ranked consequences by patients and shows the ranking of these consequences by knee specialists. The overall prioritizing of all 29 consequences after TKA by patients and knee specialists are shown in Figure 4 and Supplementary Table 1. The highest ranked item for patients was "Inability to do normal activities such as walking, cycling, swimming, heavy household chores," while "No improvement in pain during the day" was highest ranked by knee specialists. This latter item was not ranked in the top-5 of patients, whereas "No improvement in walking" was in the patients' top-5 but was not ranked in the top-5 of knee specialists. In total, four out of five items in the top-5 of both patients and knee specialists were similar; however, the ranking was different.

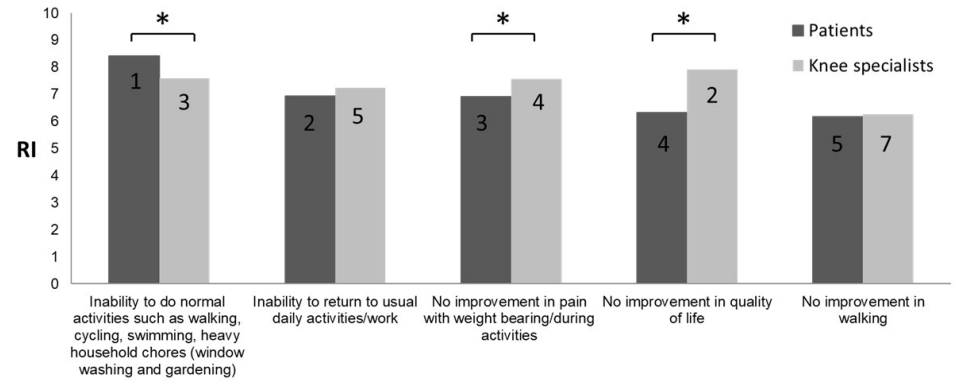


Figure 3. Relative importance scores of patients (dark-shaded bars) and knee specialists (light-shaded bars) of the top 5 most important postoperative consequences for patients after TKA as prioritized by patients. The numbers on the bars indicate the ranking. * = significant differences between groups ($P < .05$).

Agreement in ranking of RIs

A high level of agreement was found for the ranking of RIs between patients versus knee specialists (KW 0.91, *P* value .005) and satisfied patients and dissatisfied patients (KW 0.92, *P* value .004) according to Kendall's W analyses.

Differences in RIs

The top-5 items with the largest differences in RIs between patients and knee specialists are shown in Table 2. The item that was ranked higher by patients than by knee specialists was "Unable to walk outdoors with aids" and the item that was ranked higher by knee specialists than by patients was "No improvement in pain during the day." "No improvement in pain with weight bearing/during activities" was more important in dissatisfied patients compared to satisfied patients. Further exploration of differences between satisfied and dissatisfied patients are shown in Table 3.

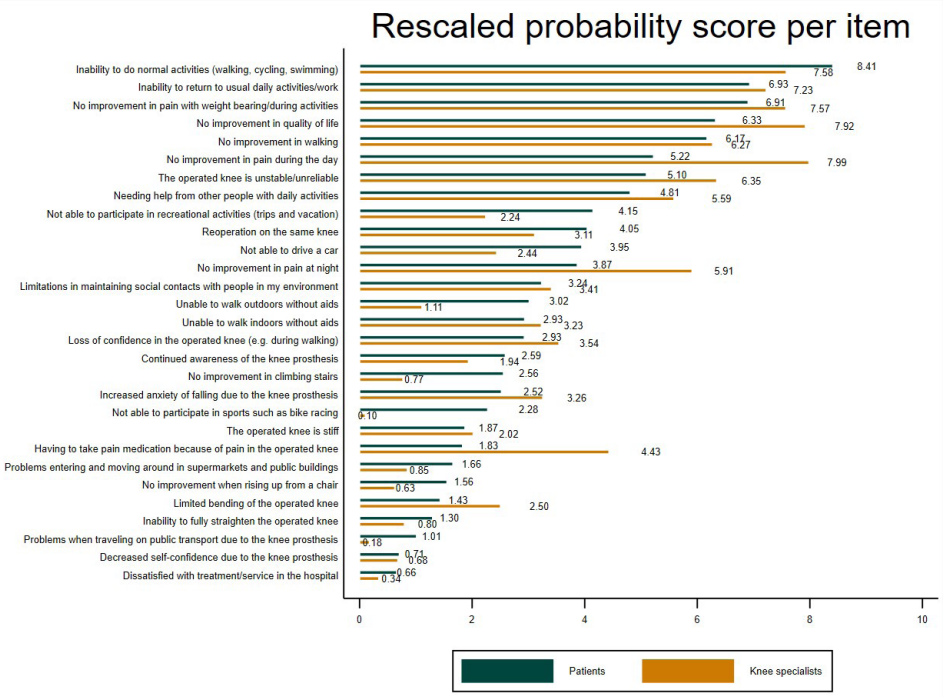


Figure 4. RIs of items of the BWS exercise in patients and knee specialists ranked by importance scored by patients

Discussion

To our knowledge, this is the first study asking both TKA patients and knee specialists to rank adverse consequences for patients after TKA that contribute to poor response. The most important finding was that the perspectives of patients and knee specialists on the importance of consequences for patients are comparable. Consequences that were highly prioritized by patients were directed towards limitations they experience during walking and performing valued daily activities. Knee specialists were well aware of consequences that contribute to poor response to TKA for patients, but ranked pain-related items higher than patients did.

Table 2. Top 5 items with the largest differences in score between patients and knee specialists[a]

Postoperative consequences	Patients (n=95)		Knee specialists (n=63)		Differences in RI
	RI	ranking	RI	ranking	
No improvement in pain during the day	5.2	6	8.0	1	2.8
Having to take pain medication because of pain in the operated knee	1.8	22	4.4	10	2.6
Not being able to participate in active sports such as bike racing	2.3	20	0.1	29	2.2
No improvement in pain at night	3.9	12	5.9	8	2.1
Not being able to participate in recreational activities such as taking trips and going on vacation	4.2	9	2.2	18	1.9
Unable to walk outdoors with aids	3.0	14	1.1	21	1.9

RI, relative importance
[a]Convenient top 5 because 2 items have an equal RI score

Table 3. Top 5 items with the largest differences in score between satisfied and dissatisfied patients[a]

Postoperative consequences	Dissatisfied patients (n=16)		Satisfied patients (n=79)		Differences in RI
	RI	ranking	RI	ranking	
Needing help from other people with daily activities	3.0	14	5.2	6	2.1
No improvement in pain with weight bearing/during activities	8.5	1	6.6	3	1.9
Not being able to drive a car independently	2.6	19	4.2	11	1.6
Inability to return to usual daily activities/work	5.8	5	7.2	2	1.4
The operated knee is stiff	3.0	15	1.7	23	1.4
Reoperation on the same knee	2.9	16	4.3	9	1.4

RI, relative importance
[a]Convenient top 5 because 3 items have an equal RI score

This study showed comparable results to previous studies that focused on asking patients what results matter the most to patients undergoing a knee or hip arthroplasty.[16,17] Relief of pain, recovery of function, and improved quality of life are the 3 outcomes ranked highest by Goodman et al.[16] Whitebird et al. identified the ability to walk without pain/discomfort, pain relief, and returning to an active lifestyle as important outcomes.[17] While pain relief itself was the general focus in both studies, many patients discussed pain in relation to specific activities like mobility and walking. This is also reflected in our results as the top-5 ranked items in patients were all related to limitations during walking and daily activities. This finding suggests that patients are willing to accept some degree of pain after TKA as long as they are able to walk and to perform (valued) activities of daily living. One previous qualitative study[7] was concordant with our results, where patients also valued pain and limitations in physical functioning as more important than a reoperation on the same knee. These patient perspectives on relevant outcomes after TKA provides important insights that should be taken into account when developing a definition for measuring poor response to TKA.

Novel findings from our study comprise the ranking of consequences in TKA patients from the perspective of knee specialists. Overall, the ranking of knee specialists and patients was comparable and therefore we can conclude that knee specialists know which consequences for patients are the most and least important contributors to poor outcome after TKA. Notably, all pain-related items were ranked higher by knee specialists than patients, except for the item "No improvement in pain with weight bearing/during activities." This might be explained by the fact that pain is a major indication to perform TKA.[18,19]

Remarkably, satisfied and dissatisfied patients agree on the overall ranking of adverse consequences after TKA. However, the items "help from other people with daily activities" and "reoperation on the same knee" were ranked higher in satisfied patients than in dissatisfied patients. This difference may be explained by the fact that satisfied patients rate a hypothetical situation and dissatisfied patients give a high ranking to consequences that they are currently experiencing.

These findings have several important clinical and research implications. The results of this study could be used during consultation; HCPs should address expectations and concerns related to walking ability and the ability to perform valued activities after TKA. Furthermore, the most important items could be incorporated in a tool for patients to elicit their personal preferences and prepare for a preoperative consult in which a TKA procedure is discussed. Moreover, it is desirable to measure these most important items in patients in order to preoperatively inform patients about the expected results. For research purposes, we can conclude that return to normal activities/daily work, daily knee functioning (limited by pain), quality of life, and walking ability are important outcome domains according to the patient perspective and should be included in a definition for measuring poor response after TKA.

The main strength of our study is that it focuses on the ranking of consequences for TKA patients contributing to poor response according to both patients and knee specialists. We included a wide range of HCPs, that were not included in previous studies. The items were taken from a previous qualitative study among both patients and HCPs. Co-working with patient research partners to develop the study documents was invaluable to ensure that the language and approaches used were accessible to patient participants. A potential limitation of a BWS exercise is that participants can only rate the items presented to them. There might be other important items that were not mentioned in the interview study due to social desirability. Also, considering the patient participant response rate of 51% there might be some selection bias. Patients were asked to complete a quite complicated task which may have prevented low literate people to participate in this study. We invited patients from a single hospital and we have no characteristics of the non-responders, therefore responders and non-responders could not be compared. The proportion satisfied and dissatisfied patients in this study corresponds to what is reported in the literature.[2-4] Although no clear guidelines for a minimal sample size for a BWS exercise are given, in our view our sample size did not allow comparisons between subgroups of HCPs. The diversity of HCPs assures a good representation of HCPs that are involved in TKA care in the Netherlands. However, because of differences in health care systems, generalizability to other countries should be taken with caution. Replication of this BWS exercise in other countries would therefore be of interest.

Conclusions

Comparable perspectives were found for patients versus knee specialists and satisfied patients versus dissatisfied patients on the importance of consequences after TKA that contribute to a poor response. Return to normal activities/daily work, daily knee functioning (limited by pain), quality of life and walking ability are important outcome domains according to the patient perspective that should be included in a definition for defining poor response after TKA. Furthermore, knee specialists should address expectations and concerns related to the inability to perform valued activities after TKA during consultation.

Acknowledgements

The authors wish to thank Saskia Susan for patient recruitment, and Wilma Peters, Manorma Dwarkasing, and Ine Willems (patient research partners) for their contribution to this study.

References

1. Batailler C, Swan J, Sappey Marinier E, Servien E, Lustig S. New Technologies in Knee Arthroplasty: Current Concepts. J Clin Med 2020;10(1)
2. DeFrance M, Scuderi G. Are 20% of Patients Actually Dissatisfied Following Total Knee Arthroplasty? A Systematic Review of the Literature. J Arthroplasty 2022;38(3):594-599.
3. Reynaud V, Verdilos A, Pereira B, Boisgard S, Costes F, Coudeyre E. Core Outcome Measurement Instruments for Clinical Trials of Total Knee Arthroplasty: A Systematic Review. J Clin Med 2020;9(8):2439.
4. Halawi MJ, Jongbloed W, Baron S, Savoy L, Williams VJ, Cote MP. Patient Dissatisfaction After Primary Total Joint Arthroplasty: The Patient Perspective. J Arthroplasty 2019;34(6):1093-1096.
5. te Molder MEM, Smolders JMH, Heesterbeek PJC, van den Ende CHM. Definitions of poor outcome after total knee arthroplasty: an inventory review. BMC Musculoskelet Disord 2020;21(1):378.
6. Price AJ, Longino D, Rees J, et al. Are pain and function better measures of outcome than revision rates after TKR in the younger patient? Knee 2010;17(3):196-9.
7. Singh JA, Mehta B, Mirza SZ, et al. When Has a Knee or Hip Replacement Failed? A Patient Perspective. J Rheumatol 2021;48(3):447-453.
8. Te Molder MEM, Vriezekolk JE, van Onsem S, Smolders JMH, Heesterbeek PJ, van den Ende CHM. Exploration of adverse consequences of TKA by patients and knee specialists: a qualitative study. (submitted);
9. Mühlbacher AC, Kaczynski A, Zweifel P, Johnson FR. Experimental measurement of preferences in health and healthcare using best-worst scaling: an overview. Health Econ Rev 2016;6(1):2.
10. Cheung KL, Wijnen BF, Hollin IL, et al. Using Best-Worst Scaling to Investigate Preferences in Health Care. Pharmacoeconomics 2016;34(12):1195-1209.
11. Hollin IL, Paskett J, Schuster ALR, Crossnohere NL, Bridges JFP. Best-Worst Scaling and the Prioritization of Objects in Health: A Systematic Review. Pharmacoeconomics 2022;40(9):883-899.
12. Fortin PR, Penrod JR, Clarke AE, et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. Arthritis Rheum 2002;46(12):3327-30.
13. Lingard EA, Katz JN, Wright EA, Sledge CB. Predicting the outcome of total knee arthroplasty. J Bone Joint Surg Am 2004;86(10):2179-86.
14. Nilsdotter AK, Toksvig-Larsen S, Roos EM. A 5 year prospective study of patient-relevant outcomes after total knee replacement. Osteoarthritis Cartilage 2009;17(5):601-6.
15. Sawtooth Software. Identifying "bad" responders: fit statistic and identifying random responders. Retrieved from: https://www.sawtoothsoftware.com/help/lighthouse-studio/manual/index.html?hid_web_maxdiff_badrespondents.html. Date last accessed: 02.03.23.
16. Goodman SM, Mehta B, Mirza SZ, et al. Patients' perspectives of outcomes after total knee and total hip arthroplasty: a nominal group study. BMC Rheumatol 2020;4:3.
17. Whitebird RR, Solberg LI, Ziegenfuss JY, et al. Personalized outcomes for hip and knee replacement: the patients point of view. J Patient Rep Outcomes 2021;5(1):116.
18. Cross WW, Saleh KJ, Wilt TJ, Kane RL. Agreement about indications for total knee arthroplasty. Clin Orthop Relat Res 2006;446:34-9.
19. Gademan MG, Hofstede SN, Vliet Vlieland TP, Nelissen RG, Marang-van de Mheen PJ. Indication criteria for total hip or knee arthroplasty in osteoarthritis: a state-of-the-science overview. BMC Musculoskelet Disord 2016;17(1):463.

Supplementary material

Supplementary table 1. Total list of consequences for patients after total knee arthroplasty (TKA) (n = 29) used in the best-worst scaling exercise, ranked by TKA patients and knee specialists, with relative importance scores (RI).

Rank	Ranking patients Adverse consequences for patients with OA	RI
1	Inability to do normal activities such as walking, cycling, swimming, heavy household chores (window washing and gardening)	8.41
2	Inability to return to usual daily activities/work	6.93
3	No improvement in pain with weight bearing/during activities	6.91
4	No improvement in quality of life	6.33
5	No improvement in walking	6.17
6	No improvement in pain during the day	5.22
7	The operated knee is unstable/unreliable	5.1
8	Needing help from other people with daily activities	4.81
9	Not being able to participate in recreational activities such as taking trips and going on vacation	4.15
10	Reoperation on the same knee	4.05
11	Not being able to drive a car independently	3.95
12	No improvement in pain at night	3.86
13	Limitations in maintaining social contacts with people in my environment (family, neighborhood, friends) due to the knee prosthesis	3.24
14	Unable to walk outdoors without aids	3.02
15	Unable to walk indoors without aids	2.93
16	Loss of confidence in the operated knee, for example during walking	2.93
17	Continued awareness of the knee prosthesis	2.59
18	No improvement in climbing stairs	2.56
19	Increased anxiety of falling due to the knee prosthesis	2.52
20	Not being able to participate in active sports such as bike racing	2.28
21	The operated knee is stiff	1.87
22	Having to take pain medication because of pain in the operated knee	1.83
23	Problems entering and moving around in supermarkets and public buildings (e.g. healthcare, government) due to the knee prosthesis	1.66
24	No improvement when rising up from a chair	1.56
25	Limited bending of the operated knee (less than a right angle such as sitting on a chair)	1.43
26	Inability to fully straighten the operated knee	1.3
27	Problems when traveling on public transport due to the knee prosthesis	1.01
28	Decreased self-confidence due to the knee prosthesis (e.g. due to misunderstanding of the situation by others or due to the scar)	0.71
29	Dissatisfaction with treatment/service in the hospital	0.66

Rank	Ranking knee specialists Adverse consequences for patients with osteoarthritis	RI
1	No improvement in pain during the day	7.99
2	No improvement in quality of life	7.92
3	Inability to do normal activities such as walking, cycling, swimming, heavy household chores (window washing and gardening)	7.58
4	No improvement in pain with weight bearing/during activities	7.57
5	Inability to return to usual daily activities/work	7.23
6	The operated knee is unstable/unreliable	6.35
7	No improvement in walking	6.27
8	No improvement in pain at night	5.91
9	Needing help from other people with daily activities	5.59
10	Having to take pain medication because of pain in the operated knee	4.43
11	Loss of confidence in the operated knee, for example during walking	3.54
12	Limitations in maintaining social contacts with people in my environment (family, neighborhood, friends) due to the knee prosthesis	3.41
13	Increased anxiety of falling due to the knee prosthesis	3.26
14	Unable to walk indoors without aids	3.23
15	Reoperation on the same knee	3.11
16	Limited bending of the operated knee (less than a right angle such as sitting on a chair)	2.50
17	Not being able to drive a car independently	2.44
18	Not being able to participate in recreational activities such as taking trips and going on vacation	2.24
19	The operated knee is stiff	2.02
20	Continued awareness of the knee prosthesis	1.94
21	Unable to walk outdoors without aids	1.11
22	Problems entering and moving around in supermarkets and public buildings (e.g. healthcare, government) due to the knee prosthesis	0.85
23	Inability to fully straighten the operated knee	0.80
24	No improvement in climbing stairs	0.77
25	Decreased self-confidence due to the knee prosthesis (e.g. due to misunderstanding of the situation by others or due to the scar)	0.68
26	No improvement when rising up from a chair	0.63
27	Dissatisfaction with treatment/service in the hospital	0.34
28	Problems when traveling on public transport due to the knee prosthesis	0.18
29	Not being able to participate in active sports such as bike racing	0.10

Chapter 8

General discussion

Introduction

To date, there is no consensus how to define poor response to total knee arthroplasty (TKA). Differences in perspectives on poor outcome after TKA by patients and knee specialists might be explained by this lack of a uniform vocabulary. Therefore, in this thesis both patients' and knee specialists' perspectives on poor response to TKA were qualitatively and quantitatively studied in various studies.

In this section, I elaborate on the main findings of this thesis and discuss methodological and other issues that emerged during the process of doing my research. I will end with implications for clinical practice, suggestions for future research directions and several conclusions.

Concept of poor response after total knee arthroplasty

The concept of poor response after TKA is best reflected by patient-reported 'treatment failure'.[1,2] Treatment failure is a novel anchor-based method in responder analysis to define those patients who find their treatment has failed.[3,4] Responder analysis is a concept where each participant in a study is classified as either a 'responder' or a 'non-responder' to the treatment.[5] The findings of this thesis demonstrate that patient (dis)satisfaction is currently used as the most common measure of treatment failure both in clinical practice and in the literature. This is not remarkable, given that (dis)satisfaction as an outcome domain is incorporated in both the International Consortium for Health Outcomes Measurement (ICHOM) core outcome domain set for hip and knee osteoarthritis as well as the Outcome Measures in Rheumatology-Osteoarthritis Research Society International (OMERACT-OARSI) core domain set for clinical trials of total joint replacement (TJR).[6,7] However, what is actually measured with 'patient (dis)satisfaction' and how is this related to other outcome domains such as pain, functioning, and quality of life? According to ICHOM, measures such as satisfaction, fulfillment of expectations, and willingness to repeat or to recommend treatment to others were not seen as true patient-reported outcome measures (PROMs).[6] ICHOM states that these measures are associated with changes in PROM scores and may reflect more how well a provider manages to set realistic expectations on outcomes.[6] The OMERACT-OARSI working group has not even described what is measured with the satisfaction outcome domain. This thesis has shown that a definition on (dis)satisfaction with the outcome of TKA for measuring poor outcome after TKA received the highest scores for feasibility (Chapter 6). The understanding of (dis)satisfaction after TKA is limited by 2 factors: (first) a lack of a clear definition of patient (dis)satisfaction, and (second) limitations in the way we measure (dis) satisfaction. The wide range of reported dissatisfaction rates can be attributed to the diversity in answering options, wording, and timeframe used across studies. It is important to acknowledge that this variability in measurement methods contributes to the challenges in accurately assessing and comparing dissatisfaction rates, and thus treatment failure, among different studies. If we continue to define treatment failure as patient dissatisfaction it is important to reach consensus on what and how (dis)satisfaction after TKA should be measured.

Patient involvement in research

In order to tailor a definition of poor response to TKA to the perspective of patients and knee specialists, patients were actively involved in the various studies described in this thesis. Patients participated as study participants and respondents in the interview study and best-worst scaling exercise (Chapter 4 and 7). In addition, three patients contributed to this thesis

as advisors and research partners. Co-working with the patient research partners for contemplation on results, and to develop study documents (e.g. various interview guides and best-worst scaling item list) was invaluable to ensure that the language and approaches used were accessible to patient participants. There is a growing trend in healthcare research to focus more on outcomes that matter to patients, and more widely on patient-centered research.[8,9] The involvement of patients not just as subjects of research but as partners in the design, assessment, and implementation of health research is recommended, and this is sometimes mandatory for grant approval.[10,11] The term 'patient research partner' (PRP) is used when referring to higher levels of engagement.[12] A PRP is someone living with the relevant disease or condition who participates as an active member on an equal basis with professional researchers, adding the benefit of his/her experiential knowledge to the research project.[12] Not all study designs in this thesis were feasible for patient involvement. For example, we deliberately chose not to include patient representatives in the international Delphi study (Chapter 6) considering the need for English language skills, as well as the complexity associated with understanding the definitions, underlying questionnaires and various approaches for the interpretation of PROMs. Instead, patient input from the interview study was incorporated into the Delphi study as new definitions were formulated based on the views of patients. In addition, a separate study on the prioritization of adverse consequences after TKA among patients was performed. For such complex study designs, it is questionable whether the PRPs eligible for involvement represent the patient group, as only patients trained in research, methodology and language are considered eligible for involvement. According to de Wit, involving PRPs regardless of their background and experience is always better than not involving patients at all.[12,14] Following this reasoning, we should have asked the PRPs if they would like to be involved as PRP in the Delphi study instead of deciding ourselves not to involve them.

A hypothetical situation

A recurring challenge in this thesis is to survey preoperative patients, satisfied patients and knee specialists about the hypothetical situation of poor outcome after TKA. As this thesis was based on cross-sectional data, it concerns situations that do not apply to these respondents or situations that could happen in the future. Hypothetical questions are based on supposition, not on real situations. They are typically used to elicit opinions and beliefs about imagined situations or conditions that do not exist. Practices involving talking in hypothetical terms and posing hypothetical questions have been documented across several settings such as cardiology[14], oncology[15] and psychiatry.[16] Evidence indicates that these questions are highly effective in encouraging patients to engage with difficult issues.[17] Whereas we used this way of questioning to generate information that we would otherwise have missed or that we would be unaware of. For example, information about preoperative expectations regarding poor outcome after TKA. It is important to mention that we also had to deal with the shortcomings of this type of questions. For instance, there were patients and knee specialists who asked for clarification during the interviews or mentioned that they found it difficult to complete the best-worst scaling study as they experienced difficulties with imagining certain situations. It is unclear whether the hypothetical situations influenced the results. Patients may have trivialized or exacerbated the situation, which could have led to an over- or underestimation of the results. To resolve this knowledge gap, a longitudinal, qualitative study is recommended to examine whether expectations and the perspective on poor outcome after TKA change over time and what influences the change or stability.

Design of the thesis

The work described in this thesis started with a structured planned workflow to gain an in-depth understanding of the concept of poor response to TKA (Figure 1). Most chapters are closely related and follow each other in sequence. Chapter 2 diverges, because the questionnaire central in this chapter, akin to numerous other patient-reported outcome measures (PROMs), primarily focuses on assessing patient improvement following TKA. The focus of the other chapters was on patients who experienced unfavorable outcomes following TKA, i.e. patients who experienced no or too little improvement after TKA. This choice was driven by the understanding that targeting patients with poor outcome offers the greatest potential for substantial improvements. The search for an in-depth understanding of the concept of poor response to TKA started with a literature review on the variety of dichotomous definitions used in the literature to describe poor outcome to TKA. Parallel to the literature review an interview study among patients and knee specialists across Belgium and the Netherlands was scheduled to identify relevant concepts to be incorporated in a definition of poor response to TKA. As part of this thesis, it was decided to form an international expert advisory group to use their expertise throughout the various stages of this research project. One of the responsibilities of this expert group was to develop new definitions based on the results of the interview study. The definitions identified in the literature review and proposed by the expert group continued as input for the performance study, which was aimed to examine the performance of a set of 'poor response to TKA' -definitions in two existing databases. The results of these studies would be transferred to the Delphi study, aimed to assess the face validity and feasibility of definitions, next to ranking the definition of poor response to TKA amongst a panel of international knee experts. It was decided in advance not to include patients in the Delphi study and therefore a best-worst scaling study among patients and the Delphi panel was added to the research project with the aim to prioritize a maximum of 20 definitions.

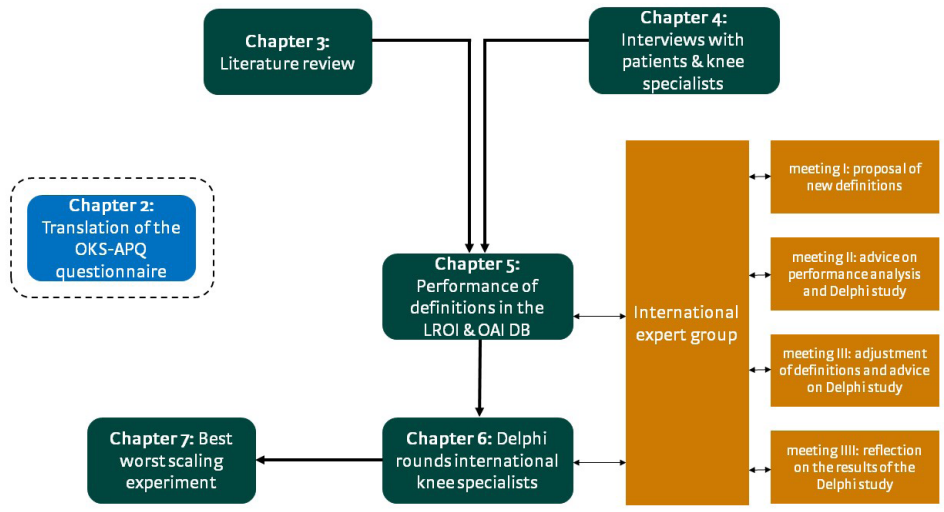


Figure 1. Schematic illustration of the workflow of this thesis, repeated from the introduction

Unforeseen modifications

The execution of the various studies largely followed the structured workflow (Figure 1), but there were some minor deviations. The COVID-19 pandemic and lock down had a significant influence. Firstly, patients in Belgium could no longer be interviewed face-to-face and we had to switch to individual online meetings with these patients. This caused a delay and may have caused missed signs of nonverbal communication. Secondly, the first expert meeting with the international expert group was initially planned as a face-to-face consensus meeting during the 3rd World Arthroplasty Conference (WAC), however, it had been converted to an online meeting as well. We utilized video conferencing instead as a second best option.

Consequently, preliminary results of the interview study were discussed with the experts, which may have prevented other new definitions from emerging. For instance, results of the interview study showed that the arduous recovery process was important for patients, but this theme was not included in the preliminary results. Nevertheless, the influence of the use of preliminary results in the follow-up studies seems limited since definitions about the recovery process did not appear in the literature review and knee experts themselves did not propose them in the Delphi study. In addition, the recovery process will probably play a role for patients in a definition on (dis)satisfaction. Another limitation of the performance study was that, other than anticipated, a smaller selection of definitions could be tested and compared due to the limited number of PROMs available in both datasets. For example, different anchors for the assessment of discriminative accuracy had to be used because no comparable generic PROMs were available in both sets. Furthermore, the Dutch Arthroplasty register (LROI) had preoperative and 12-months postoperative data available, while the data in the Osteoarthritis Initiative (OAI) dataset was time varying. We therefore set minimum and maximum time windows for pre- and postoperative data assessment, but these windows differ (in a range) per patient and are also different with the data structure of the LROI. Prospective, longitudinal research is needed to make better comparisons of definitions of poor response to TKA within data sets.

Qualitative and quantitative methods

Qualitative research explores and provides deeper insights into real-world problems.[18] Instead of dealing with numerical data and statistics, qualitative research helps to generate hypotheses, and complements quantitative research in healthcare research.[19,20] According to Green and Thorogood, the aim of qualitative research is to describe and understand social phenomena in terms of the meanings people give to them.[21] It answers the "hows and whys" instead of "how many or how much". Although qualitative research is increasingly recognized and valued in the field of health sciences, its status remains controversial.[22,23] Researchers unfamiliar with its methodologies sometimes perceive qualitative research as inferior to quantitative research, criticizing it for issues of scientific rigor, trustworthiness and generalizability of study results.[22,24] The qualitative research in this thesis was reported according to the Standards for Reporting of Qualitative Research (SRQR) checklist to ensure complete and transparent reporting.[25]

This thesis choose for such a qualitative design, followed by a quantitative design to identify adverse consequences of TKA as perceived by patients and knee specialists. A strength of this thesis is that the perspectives of patients' and knee specialists' on poor outcome to TKA were explored through a combination of qualitative and quantitative studies. An important and

unexpected finding of our qualitative study was the impact of the arduous recovery process on patients; patients were eager to share their experiences. Our research does not allow inferences about the impact of these experiences on how patients value the ultimate results of their TKA. To future explore this new topic, quantitative research is needed.

Quantitative research methods were subsequently used to further prioritize and verify the qualitative findings by analysis of large databases and a best-worst scaling (BWS) exercise. A BWS exercise with hierarchical Bayesian analysis was used to determine the relative importance of adverse consequences of TKA as perceived by patients in the qualitative study (Chapter 7). BWS allowed us to prioritize the qualitative data from a patients' perspective and resulted in valuable information from the perspective of patients themselves and from knee specialists. The preparation of the item list for the BWS study was challenging. The procedure of extracting a representative and clear set of items from the large number of statements that were obtained in 40 interviews with patients and knee specialists was not previously described in the literature. Therefore, multiple researchers were involved in the selection and formulation of the items. Also, three PRPs were involved in this study and contributed in discussing the wording of the items, refined the initial list to a final list of items and checked the final list for its completeness and comprehensiveness. Despite the input of the PRPs, it was noticed that patients and knee specialists found it difficult to prioritize the items in the BWS study. As mentioned before, this may be explained by the fact that satisfied patients and knee specialists had to rate a hypothetical situation.

Implications

The aim of this thesis was to identify a definition for the assessment of poor response after TKA. Although no ultimate quantitative definition for measuring poor response to TKA was identified, this thesis did contribute to an in-depth understanding of the outcome domains and interpretations (e.g. type of threshold, moment of follow-up, and perspectives) that should be incorporated in a definition to identify patients as poor responders after TKA. Having completed all the studies described in this thesis, the need for a good definition has been reaffirmed and is still there. This is reflected by the findings of this thesis, as well as by the great interest and willingness to participate among patients and knee specialists throughout the research process. The high number of definitions identified in the review of the literature (Chapter 3) indicates international efforts to assess the proportion of patients with poor outcome. To date, however, it has proven ineffective because many different approaches are used, making it impossible to properly compare the number of poor responders across hospitals, and between countries. (Inter)national consensus on a definition is a prerequisite to allow benchmarking across (inter)national institutions. To achieve this more research is needed. As previously discussed the performance of the six prioritized definitions (Chapter 5) should be compared in a large prospective multicenter cohort study. Patient (dis)satisfaction after TKA assessed in a standardized way could be used as the anchor for this study. Then, to create support, the results should be discussed in an international forum, for instance the OMERACT.

Moreover, the identified definitions can have direct clinical implications: as our results could guide single hospitals to choose and implement a definition of poor response to TKA in routine clinical practice for quality improvement and transparency. With that information variations in care over time and among health care providers within hospitals can be determined and

8

discussed. Learning from colleagues, through evaluation of the process of care and working together could potentially provide clearer insights for improvement than monitoring based on PROMs alone. For instance, the interview study (Chapter 4) showed that the arduous recovery process was important to patients and that is something difficult to monitor with PROMs and better to discuss with colleagues.

The definitions of poor response after TKA identified in this thesis can also be used to examine the predictive value of preoperative, modifiable factors in research and clinical practice to identify patients who are at risk of poor outcome after TKA. The most common sociodemographic, preoperative and postoperative factors for dissatisfaction have already been identified in a previous literature review.[26] For instance, psychological factors such as depression, anxiety, and pain catastrophizing were identified as common preoperative factors for dissatisfaction. Although it is the question whether the same factors will influence poor outcome after TKA, it is advisable to investigate the effect of modifying these factors. For example, this can be done by training psychological flexibility or treating pain reduction preoperatively to increase preoperative patient optimization.[26] Our results suggest that also factors related to hospitalization and the recovery process should be examined for their predictive value of poor response. In line with this, it is conceivable that interventions to better guide patients pre- and posthospitalization could improve their appreciation of the result of TKA.

The overview and prioritization of adverse consequences to TKA presented in this thesis can be a starting point for improvements in personalized care of individual patients in clinical practice. The most important adverse consequences identified by patients can, for instance, be incorporated in a tool for patients to elicit their personal preferences and prepare for a preoperative consult in which TKA treatment is discussed.

An eHealth tool to prepare for an orthopedic consultation or a decision aid supports patients and their orthopedic surgeons in having the right conversation and making well-thought treatment decisions. This will facilitate and improve personalized care in clinical practice. Over the past years, various decision aids and option grids were developed to support in TKA choices.[28-30] However, these decision aids did not include aspects regarding important adverse consequences of TKA identified by patients in this thesis. Therefore, the most important adverse consequences identified by patients (Chapter 7) could be incorporated into a tool where patients can tick or prioritize consequences they find relevant to discuss with their orthopedic surgeon. Such a tool will assist patients to explicitly state their personal preferences, personal goals and expectations. The orthopedic surgeon can use the tool as a conversation guide for discussing the relative probability that the patient would be able to accomplish each of their stated goals. In case of a discrepancy between what the patient expects and what orthopedic surgeons know TKA can deliver, the orthopedic should explain how realistic the patients' expectations are[31]. When it is possible to identify patients who are at risk of poor outcome after TKA, the contributing factors can be identified, modified and tools deployed to reduce the number of poor responders after TKA, this will undoubtedly improve the quality of orthopedic care.

Future research directions

As discussed in the previous paragraphs the research described in this thesis reveals several issues that need further investigation. Assuming unrestricted time, money and patients, two types of prospective longitudinal research are recommended. First, a prospective longitudinal, qualitative study into the perspective and expectations on poor outcome after TKA could provide useful information on whether a shift in perspective and outcome expectations occurs after patients personally experience a TKA treatment. This information is relevant as, for instance, it can show whether the perspective on poor outcome after TKA changes over time and can lead to the advice to develop specific definitions for specific follow-up moments during TKA recovery. Second, another prospective, longitudinal study would be of interest to study the pre- and postoperative measures for the identified important outcome domains (pain, daily knee functioning and (dis)satisfaction) in each patient to evaluate and compare the performance of the six prioritized definitions. Also, prognostic factors can be explored to study their contribution to poor response after TKA. That is necessary to identify and modify contributing factors in order to reduce the number of poor responders after TKA. Finally, quantitative research on the impact of the arduous recovery process on patients, identified in the interview study would be of interest. Aspects associated with the difficult process of adapting to the prosthesis need to be addressed during shared decision making based on the results of quantitative research.

Conclusions

This thesis contributed to the understanding of the concept of poor response after TKA and provided insight in how it can be best measured according to the perspective of patients and knee specialists. This thesis shows that patients and knee specialists have comparable perspectives on poor outcome after TKA. Pain, daily knee functioning and (dis)satisfaction are identified as important outcome domains according to both perspectives and should be included in a definition for defining poor response after TKA. Finally, patient-oriented definitions are preferred over physician-oriented definitions. The work in this thesis adds to the knowledge about the unhappy patient after TKA and has identified some issues that require further exploration to reduce the number of poor responders after TKA.

8

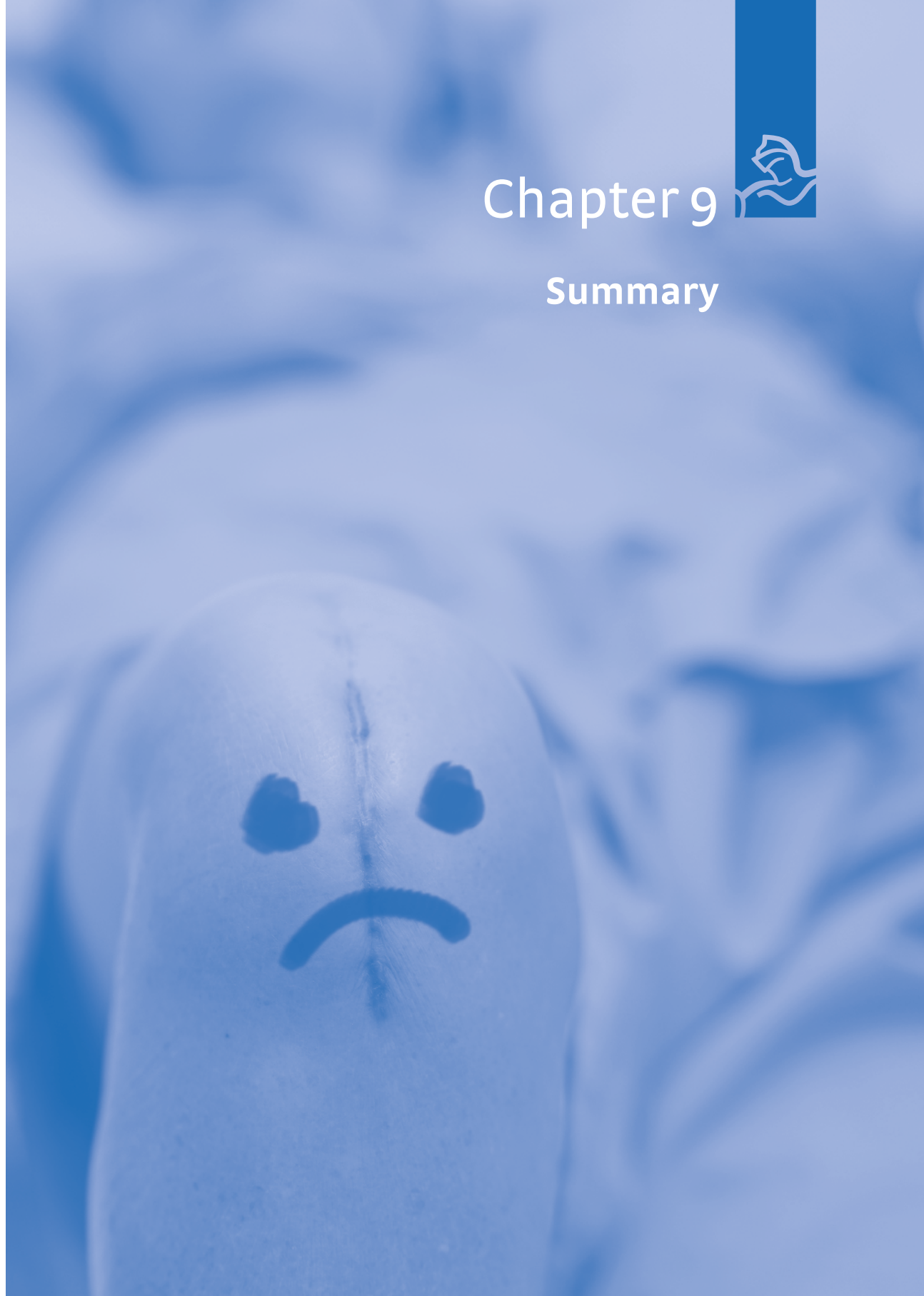
References

1. Ingelsrud LH, Granan LP, Terwee CB, Engebretsen L, Roos EM. Proportion of Patients Reporting Acceptable Symptoms or Treatment Failure and Their Associated KOOS Values at 6 to 24 Months After Anterior Cruciate Ligament Reconstruction: A Study From the Norwegian Knee Ligament Registry. The American journal of sports medicine 2015;43(8):1902-7.
2. Ingelsrud LH, Terwee CB, Terluin B, et al. Meaningful Change Scores in the Knee Injury and Osteoarthritis Outcome Score in Patients Undergoing Anterior Cruciate Ligament Reconstruction. The American journal of sports medicine 2018;46(5):1120-1128.
3. Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomised trials. Bmj 1998;316(7132):690-3.
4. Tubach F, Ravaud P, Beaton D, et al. Minimal clinically important improvement and patient acceptable symptom state for subjective outcome measures in rheumatic disorders. J Rheumatol 2007;34(5):1188-93.
5. Roos EM, Boyle E, Frobell RB, Lohmander LS, Ingelsrud LH. It is good to feel better, but better to feel good: whether a patient finds treatment 'successful' or not depends on the questions researchers ask. Br J Sports Med 2019;53(23):1474-1478.
6. Rolfson O, Wissig S, van Maasakkers L, et al. Defining an International Standard Set of Outcome Measures for Patients With Hip or Knee Osteoarthritis: Consensus of the International Consortium for Health Outcomes Measurement Hip and Knee Osteoarthritis Working Group. Arthritis care & research 2016;68(11):1631-1639.
7. Singh JA, Dowsey MM, Dohm M, et al. Achieving Consensus on Total Joint Replacement Trial Outcome Reporting Using the OMERACT Filter: Endorsement of the Final Core Domain Set for Total Hip and Total Knee Replacement Trials for Endstage Arthritis. J Rheumatol 2017;44(11):1723-1726.
8. Speight J, Barendse SM. FDA guidance on patient reported outcomes. Bmj 2010;340:c2921.
9. Staniszewska S, Haywood KL, Brett J, Tutton L. Patient and public involvement in patient-reported outcome measures: evolution not revolution. Patient 2012;5(2):79-87.
10. Brett J, Staniszewska S, Mockford C, et al. A systematic review of the impact of patient and public involvement on service users, researchers and communities. Patient 2014;7(4):387-95.
11. Wright MT, Roche B, von Unger H, Block M, Gardner B. A call for an international collaboration on participatory research for health. Health Promot Int 2010;25(1):115-22.
12. de Wit MP, Kvien TK, Gossec L. Patient participation as an integral part of patient-reported outcomes development ensures the representation of the patient voice: a case study from the field of rheumatology. RMD Open 2015;1(1):e000129.
13. Cheung PP, de Wit M, Bingham CO, 3rd, et al. Recommendations for the Involvement of Patient Research Partners (PRP) in OMERACT Working Groups. A Report from the OMERACT 2014 Working Group on PRP. J Rheumatol 2016;43(1):187-93.
14. de Wit M, Kirwan JR, Tugwell P, et al. Successful Stepwise Development of Patient Research Partnership: 14 Years' Experience of Actions and Consequences in Outcome Measures in Rheumatology (OMERACT). Patient 2017;10(2):141-152.
15. Ahluwalia SC, Levin JR, Lorenz KA, Gordon HS. "There's no cure for this condition": how physicians discuss advance care planning in heart failure. Patient education and counseling 2013;91(2):200-5.
16. Tulsky JA, Fischer GS, Rose MR, Arnold RM. Opening the black box: how do physicians communicate about advance directives? Ann Intern Med 1998;129(6):441-9.
17. Peräkylä A, Bor R. Interactional problems of addressing 'dreaded issues' in HIV-counselling. AIDS Care 1990;2(4):325-38.
18. Parry R, Land V, Seymour J. How to communicate with patients about future illness progression and end of life: a systematic review. BMJ Support Palliat Care 2014;4(4):331-41.

19. Moser A, Korstjens I. Series: Practical guidance to qualitative research. Part 1: Introduction. Eur J Gen Pract 2017;23(1):271-273.
20. Aspers P, Corte U. What is Qualitative in Qualitative Research. Qual Sociol 2019;42(2):139-160.
21. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks: Saga Publications; 2017.
22. Green J, Thorogood N. Qualitative methods for health research. 3rd ed. ed. SAGE; 2014.
23. Daly J, Willis K, Small R, et al. A hierarchy of evidence for assessing qualitative health research. J Clin Epidemiol 2007;60(1):43-9.
24. Mays N, Pope C. Assessing quality in qualitative research. Bmj 2000;320(7226):50-2.
25. Nowell LS, Norris JM, White DE, Moules NJ. Thematic Analysis: Striving to meet the thrustworthiness criteria. International Journal of Qualitative Methods 2017;16(1):1-13.
26. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014;89(9):1245-51.
27. DeFrance M, Scuderi G. Are 20% of Patients Actually Dissatisfied Following Total Knee Arthroplasty? A Systematic Review of the Literature. J Arthroplasty 2022;
28. Légaré F, Ratté S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals' perceptions. Patient education and counseling 2008;73(3):526-35.
29. Jayadev C, Khan T, Coulter A, Beard DJ, Price AJ. Patient decision aids in knee replacement surgery. Knee 2012;19(6):746-50.
30. Jayakumar P, Moore MG, Furlough KA, et al. Comparison of an Artificial Intelligence-Enabled Patient Decision Aid vs Educational Material on Decision Quality, Shared Decision-Making, Patient Experience, and Functional Outcomes in Adults With Knee Osteoarthritis: A Randomized Clinical Trial. JAMA Netw Open 2021;4(2):e2037107.
31. Patiëntenfederatie Nederland, Federatie Medisch Specialisten. Consultkaart. 2023. Retrieved from: http://consultkaart.nl/. Date last accessed: 17-07-2023.
32. Dunbar MJ, Richardson G, Robertsson O. I can't get no satisfaction after my total knee replacement: rhymes and reasons. Bone Joint J 2013;95-b(11 Suppl A):148-52.

Chapter 9

Summary



Summary

Total knee arthroplasty (TKA) is widely recognized as an effective and successful end-stage surgical procedure for relieving chronic knee pain and functional disability in patients with osteoarthritis, based on results from surgeon-based outcome tools and survivorship analysis. However, it has emerged, through the use of patient-reported outcome measures (PROMs), that this is not a true representation of the experience of TKA for all patients because there exists a considerable number of 'unhappy patients' not being satisfied with the outcome. The aim of the research presented in this thesis was to contribute to the conceptualization of poor response after TKA, in order to identify the 'unhappy patients after TKA'. In the next paragraphs the results of this thesis are summarized and main findings were formulated.

The aim of the study described in **Chapter 2** was to translate the Oxford Knee Score – Activity and Participation (OKS-APQ) questionnaire into Dutch, and to evaluate its measurement properties in pre- and postoperative TKA patients. Translation and adaptation of the questionnaire was performed according to the forward-backward translation multi step approach. Floor and ceiling effects, structural validity, construct validity, internal consistency and test-retest reliability were evaluated using the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) quality criteria. This was assessed in 131 patients, 72 preoperative and 59 postoperative to TKA. Confirmatory factor analyses and internal consistency were good, and construct validity was supported. Preliminary findings suggest that the Dutch version of the OKS-APQ is reliable and valid for a postoperative TKA patient group, however seems less suitable in a preoperative patient group because of floor effects and lower test-retest reliability.

Main finding I: The Dutch version of the OKS-APQ can be used alongside the OKS to discriminate among levels of activity and participation in postoperative TKA patients.

Clinical research on TKA outcomes is prevalent in the literature. However, results of measurement instruments and definitions can be difficult to compare and contrast, because many different instruments and definitions are used to evaluate TKA outcomes and identify patients with an unfavorable course. In addition, these studies sometimes have poor methodological and reporting quality. Therefore the aim of the study described in **Chapter 3** was to review the literature and summarize dichotomous definitions of poor response to TKA based on a systematic search in three scientific databases. In total, 43 studies included 57 definitions of poor outcome after TKA. Ultimately, 47 different dichotomous definitions of poor outcome after primary TKA were included in the study. These 47 different definitions varied in nature and number of outcome domains involved, the type of response and the magnitude of change. A total of eight different dimensions were used in identified definitions of poor outcome: pain, function, physical functioning, health-related quality of life, patient satisfaction, anxiety, depression and patient global assessment. The absolute cut-off value was the most common type of threshold, with large variety in value and timing of follow-up.

Main finding II: Various dichotomous definitions comprising one or more outcome dimensions are used to quantify the proportion of patients with a poor outcome after TKA. This large variety of definitions impedes the comparisons of poor response to TKA over time, across hospitals and countries, and to identify patients at risk.

A successful outcome according to the knee specialist is not a guarantee for treatment success as perceived by patients. It is therefore important to get insight into adverse consequences of TKA by patients and knee specialists. In **Chapter 4** we described a qualitative interview study, on (adverse) outcome expectations and experiences of patients and knee specialists and differences in views between patients and knee specialists. The results were categorized into four themes: lingering pain, stagnating mobility, complications and revision surgery, and getting used to the prosthetic knee. Patients in our study were focused on the arduous process of getting used to the prosthesis, lingering pain, awareness of the artificial knee and limitations they experience during valued and daily activities. At the same time, knee specialists put emphasis on surgical failure, unexplained pain, limited walking ability and impairments that limit patients' physical functioning. These differences in perspectives between patients and knee specialists might contribute to discrepancies in their perception of poor response to TKA.

Main finding III: Knee specialists put more emphasis on surgical failure, unexplained pain, limited walking ability and impairments that limit patients' physical functioning, while patients' experiences were more focused on the arduous process of getting used to the prosthesis, lingering pain, awareness of the artificial knee and limitations they experience during valued and daily activities.

Insight in the prevalence of poor responders, overlap of definitions and discriminative accuracy in databases with large patient samples and large sets of PROM scores, provide the opportunity to compare the performance of different definitions of poor response to TKA. This information can serve as input for further research to reach consensus on the definition for poor response. Therefore, the aim of the study described in **Chapter 5** was to compare the prevalence, overlap and discriminative accuracy of 15 different definitions of poor response to TKA using two large databases. The definitions were retrieved from the literature review (Chapter 3) or newly composed by the expert advisory group based on the results of the qualitative study (Chapter 4). Data of patients one year after primary TKA from the Dutch Arthroplasty Register (LROI) (n=12,275) and the Osteoarthritis Initiative (OAI) database (n=204) were used to examine the prevalence, overlap (estimated by Cohen's kappa) and discriminative accuracy (sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and Youden index) of 15 different definitions tested at 12 months follow-up, or measuring a change score from preoperative to 12 months postoperative. The prevalence of poor responders was discordant (varied from 5.9% to 34.7% poor responders) across the different definitions, and the majority of pairs of definitions showed only a 'fair' or 'moderate' agreement. The moderate overlap found between the definitions suggests that they are not interchangeable and likely measure different aspects of poor outcome after TKA.

Main finding IV: None of the definitions we examined adequately classified poor responders of TKA. In contrast, absence of poor response to TKA could be classified with confidence. In line with this, a lack of overlap between different definitions of poor response to TKA was observed.

Various definitions are used to assess unfavorable outcomes following TKA. However, none of these definitions has demonstrated optimal performance. The establishment of an internationally accepted definition for evaluating poor response to TKA remains elusive, and it is conceivable that not all definitions hold equal applicability. Therefore, the study described in **Chapter 6** assessed the view of an international panel of knee experts regarding face validity, feasibility and relevance of definitions for measuring poor response to TKA. An international, three-round, online modified Delphi study was conducted with 69 panelists from 23 countries. A definition on (dis)satisfaction with the outcome of TKA obtained the highest scores for face validity and feasibility, and a definition reflecting change since pre-operative status from the perception of patients were highest prioritized.

Main finding V: To characterize "the unhappy patient after TKA", seventeen potential definitions of poor response to TKA were identified, of which six justify further study and potential implementation in quality assessment studies. All six definitions are patient centered.

The variety of potential adverse consequences of TKA that could contribute to a poor response were identified in Chapter 4. However, adverse consequences of TKA for patients were not previously ranked by knee specialists and compared to the ranking by patients. The study in **Chapter 7** described a Best-Worst scaling exercise to prioritize adverse consequences of TKA for patients that contribute to a poor response, according to the perspective of patients and knee specialists. 95 postoperative patients and 63 knee specialists prioritized a set of 29 consequences, derived from the previous qualitative study (Chapter 4). The highest-ranked consequence for patients was "Inability to do normal activities such as walking, cycling, swimming and heavy household chores" while knee specialists ranked "No improvement in pain during the day" highest. Comparable perspectives were found for patients versus knee specialists on the importance of adverse consequences after TKA. However, when looking in more detail, differences in ranking of specific subitems suggest that patients place slightly more importance on the inability to perform valued activities, while knee specialists prioritize lack of pain relief to a higher degree.

Main finding VI: Knee specialists are well aware of consequences that contribute to poor response to TKA for patients, but ranked pain-related items higher than patients did. Patients place slightly more importance on the inability to perform valued activities.

The work in this thesis adds to the knowledge about 'the unhappy patient after TKA' through a deeper understanding of the concept of poor response after TKA, and provided insight in how it can be best measured according to the perspective of patients and knee specialists. Issues were identified that require further exploration. Ultimately, if we can identify these patients better, this will facilitate quality improvement, leading to a reduction of the number of poor responders after TKA and an increase in the number of happy patients.

Chapter 10

Nederlandse samenvatting

Nederlandse samenvatting

Voor veel mensen met ernstige klachten door artrose kan een knieprothese helpen. Meestal gaat het goed, maar ongeveer één op de vijf mensen is niet blij met de knieprothese. Zij hebben nog veel pijn, kunnen niet goed lopen of voelen zich onstabiel in de knie. Tot nu toe is er geen overeenstemming over hoe we moeten bepalen wanneer een knievervangende operatie niet geslaagd is. Mensen met knieartrose of een knieprothese en kniespecialisten kunnen hierover verschillen van mening.

Dit onderzoek richtte zich daarom op het begrijpen en duidelijk maken van wat we bedoelen met een "niet-geslaagde knievervangende operatie", ook wel "slechte uitkomst" genoemd.

In **hoofdstuk 2** hebben we een Engelstalige vragenlijst over problemen in het dagelijks functioneren en bij het uitvoeren van activiteiten vertaald naar het Nederlands en uitgetest bij mensen met knieartrose of een knieprothese. Hierbij werd gekeken of de vragenlijst meet wat het zou moeten meten en of mensen consistente antwoorden op deze vragenlijst geven. De vertaalde vragenlijst bleek goed te werken bij mensen die al een knieprothese hadden, maar minder goed bij mensen die er nog geen hadden.

Hoofdstuk 3 ging over een zoektocht in de wetenschappelijke literatuur naar verschillende manieren waarop we kunnen bepalen of een knievervangende operatie niet geslaagd is. We vonden in totaal 57 manieren om een niet-geslaagde knievervangende operatie te definiëren. Deze definities verschilden veel, bijvoorbeeld in welke problemen ze meten, hoe ze meten en wat ze als niet-geslaagd beschouwen. Deze veelheid van definities maakt het lastig om resultaten van knievervangende operaties over de tijd, tussen ziekenhuizen en tussen landen te vergelijken.

Mensen met knieartrose of een knieprothese en kniespecialisten zijn geïnterviewd om te begrijpen hoe zij denken over een niet-geslaagde knievervangende operatie; hiervan wordt verslag gedaan in **hoofdstuk 4**. Verschillen in perspectieven kwamen naar voren, waarbij mensen met een knieprothese gefocust waren op het moeizame proces van wennen aan de knieprothese, aanhoudende pijn, het zich bewust zijn van de knieprothese en beperkingen die zij ervaren tijdens voor hun belangrijke dagelijkse activiteiten. Kniespecialisten daarentegen legden meer de nadruk op chirurgisch falen, onverklaarbare pijn, een beperkt loopvermogen en beperkingen in het dagelijks functioneren.

In **hoofdstuk 5** hebben we 15 verschillende definities voor een niet-geslaagde knievervangende operatie met elkaar vergeleken. Gegevens van een Nederlandse en internationale database werden daarvoor gebruikt. Over het algemeen was het moeilijk om een niet-geslaagde knievervangende operatie goed vast te stellen met deze definities, maar ze werkten wel goed om uit te sluiten welke mensen na een knievervangende operatie geen problemen hadden. De matige overlap die werd gevonden tussen definities suggereert dat definities niet uitwisselbaar zijn en waarschijnlijk verschillende aspecten van een niet-geslaagde knievervangende operatie meten.

Hoofdstuk 6 beschrijft een studie met een internationaal panel van 51 knie experts uit 23 landen die naar de definities keek en beoordeelde welke het beste waren. Definities die over ontevredenheid gingen en definities die keken naar veranderingen in vergelijking met voor de operatie werden het meest gewaardeerd. Zes definities werden aanbevolen voor verder onderzoek.

Hoofdstuk 7 ging over het vergelijken van meningen van mensen met een knieprothese en kniespecialisten over problemen na een knievervangende operatie. Deelnemers werden gevraagd de ergste problemen te kiezen uit een lijst van 29 mogelijke negatieve gevolgen. Er waren verschillen, maar ook overeenkomsten in wat ze belangrijk vonden. Het hoogst gerangschikte nadelige gevolg voor mensen met een knieprothese was “Geen normale activiteiten kunnen uitvoeren zoals wandelen, fietsen, zwemmen en zwaar huishoudelijk werk”, voor kniespecialisten was dit “Geen verbetering in pijn gedurende de dag”.

Dit onderzoek heeft geholpen om te begrijpen wat we bedoelen met een niet-geslaagde knievervangende operatie en hoe we dit het beste kunnen meten volgens mensen met knieartrose of een knieprothese en kniespecialisten. Pijn, dagelijkse beweging en tevredenheid zijn belangrijk volgens beide groepen en zouden moeten worden opgenomen in een definitie voor een niet-geslaagde knievervangende operatie.

Dankwoord

Dankwoord

Eindelijk, mijn proefschrift is klaar! De afgelopen jaren stonden voor mij in het teken van het mogen krijgen van twee prachtige dochters en de bevalling van dit proefschrift. Het was hard werken, maar vooral erg leerzaam en waardevol. Dit proefschrift als eindresultaat was niet mogelijk geweest zonder de inzet en hulp van een heleboel patiënten, zorgverleners, collega's en andere mensen om mij heen. Iedereen die heeft bijgedragen aan het tot stand komen van dit proefschrift wil ik dan ook hartelijk bedanken, en een aantal mensen in het bijzonder.

Dr. Heesterbeek, beste Petra, ik heb bewondering voor je betrokkenheid in mijn werk, privé en alles daar omheen. Je hebt me de mogelijkheid gegeven om dit proefschrift te schrijven. Bedankt voor het gestelde vertrouwen in mij en voor je altijd secure en kritische blik. Dit heeft mijn proefschrift en mijn denken zoveel beter, scherper en analytischer gemaakt. Dank daarnaast voor al je gezelligheid en interesse. Ik kon altijd bij je terecht voor vragen, advies, of om bij te kletsen. Ik kijk met veel plezier terug op de vele mooie momenten die we samen hebben meegemaakt, van mijn allereerste congres in Londen tot aan het cementeren van een knieprothese en de stormbaan in Deventer.

Dr. van den Ende, beste Els, als dagelijks begeleider heb ik ontzettend veel van je geleerd. Naast de inhoudelijke discussies was er altijd ruimte om de ups en downs te bespreken die gepaard gaan met het promoveren. Ik ben je in het bijzonder dankbaar voor de intensieve en fijne begeleiding gedurende de laatste periode van dit traject. Je hebt er altijd voor gezorgd dat ik vertrouwen bleef houden om dit promotietraject tot een goed einde te brengen. Bedankt voor de fijne samenwerking en je betrokkenheid, het ga je goed en ga lekker genieten van je welverdiende pensioen!

Dr. Smolders, beste José, dankzij jouw klinische en kritische blik op mijn stukken werd ik keer op keer aan het denken gezet. Met als uiteindelijk resultaat betere artikelen en een stevige basis voor mijn verdediging. Veel dank voor het delen van jouw ervaring en kennis. Ik heb veel waarde gehecht aan jouw betrokkenheid bij mijn ontwikkeling op werk en privé gebied, zelfs vanuit Ottawa.

Prof. Dr. de Kleuver, beste Marinus, dank voor je betrokkenheid als promotor bij het schrijven van mijn proeve van bekwaamheid tot het zelfstandig beoefenen van de wetenschap. Je waakte zowel over de voortgang van het promotietraject als over mijn persoonlijke ontwikkeling en welzijn. Bedankt voor het delen van je enthousiasme en het in mij gestelde vertrouwen. Leuk dat we aan het einde van mijn promotietraject nog wat intensiever contact met elkaar hebben gehad.

Prof. Dr. Philip van der Wees, beste Philip, bedankt voor je ondersteuning als mentor. Het was prettig om je te spreken en samen te reflecteren op mijn promotietraject.

De leden van de manuscriptcommissie bestaande uit Prof. dr. K.C.P. Vissers, Prof. dr. H.J. Schers en Prof. dr. H. Vandenneucker wil ik hartelijk danken voor het beoordelen van het manuscript en deelname in de oppositie.

Een woord van dank ook aan de LROI voor de financiering van de studies die geleid hebben tot dit proefschrift. Liza, bedankt voor de fijne samenwerking.

Special thanks to the expert advisory group. Unfortunately, the impact of the COVID-19 pandemic and lockdown prevented us from having a face-to-face meeting, but your expertise significantly contributed to various stages of this research project.

Een artikel schrijf je nooit alleen. Veel dank gaat uit naar al mijn medeauteurs. Jullie klinische input, specifieke kennis en interpretatie van de data tilden de artikelen naar een hoger niveau. Joke, Lise, Stefaan, Michelle, Ola, Jasvinder, Liza en Menno, nogmaals dank voor alle hulp bij het schrijfwerk. Een aantal medeauteurs wil ik nog persoonlijk bedanken. Joke, ontzettend bedankt voor de brainstormsessies en alle hulp bij het interview onderzoek, de fijne samenwerking en alle keren dat ik gewoon bij je binnen kon lopen met vragen. Lise, dank voor je betrokkenheid en leuk dat we samen een artikel hebben mogen schrijven. Stefaan, bedankt voor het meedenken, je klinische blik en natuurlijk het opzetten van de kwalitatieve studie. Dankzij jou konden we er een internationale studie van maken.

Beste Manorma en Wilma, bedankt voor jullie betrokkenheid en input als onderzoekspartners bij de verschillende studies. Het was waardevol en gezellig om met jullie over de onderzoeken te praten.

Uiteraard was dit proefschrift er niet geweest zonder studiedeelnemers. Graag wil ik alle patiënten en zorgprofessionals bedanken die aan de verschillende studies hebben deelgenomen.

Na 8 jaar te hebben gewerkt bij de Sint Maartenskliniek wil ik alle lieve en leuke (oud-) collega's van de Research afdeling bedanken. Ik denk met veel plezier terug aan de gezellige tijd, lunchwandelingen, borrels, cake-van-de-week momenten, kerstdiners en activiteiten in het W-gebouw. Ik heb het ontzettend naar mijn zin gehad. Tim (mede-Achterhoeker) en Bart, bedankt voor jullie humor, sarcasme en de luchtige gesprekken. Verder wil ik mijn (oud-) OrthoResearch collega's bedanken. Bij ieder van jullie stond de deur altijd open voor 'een korte vraag' maar ook om gewoon even bij te kletsen. Leuk dat de meesten van jullie twee keer naar Zieuwent-Zuid zijn afgereisd om te komen kraamschudden. Dit heb ik erg kunnen waarderen.

Mijn SMK carrière begon ik als onderzoeksverpleegkundige bij Saskia en Jolanda. Deze lieve dames wil ik in het bijzonder bedanken voor hun warme welkom, vertrouwen, steun en alles wat ze voor me gedaan hebben. Jullie lieten me geloven dat ik dit promotietraject tot een goed einde zou brengen. Samen hebben we veel gedeeld en jullie waren er altijd voor me! Saskia, mijn gehele carrière in de SMK heb je dicht bij me gestaan en daarom ben ik trots en blij dat je naast me staat als paranimf.

Miranda, onder het genot van een kopje koffie heb ik met veel plezier met je samengewerkt aan de SLIPS studie en fijn met je kunnen sparren over van alles en nog wat. Katrijn, bedankt voor je inzichten, het relativeren en niet te vergeten je humor en enthousiasme! Je pomodoro techniek komt nog dagelijks van pas bij een gezin met twee jonge kinderen;) En natuurlijk Ramon, Ilse en Kelly, bedankt voor jullie interesse, hulp en de leuke tijd!

Lieve Maartje, als kamergenootjes hebben we veel lief en leed gedeeld. Met veel plezier denk ik terug aan onze gesprekken over het weekend, 'de mini', auto's, top 2000, wielrennen en zoveel meer. Bedankt ook voor je zorg voor onze planten en je klusadviezen. Fijn dat we als oud-collega's nog contact hebben en betrokken zijn bij elkaars werk en privé ontwikkelingen. Nog even en dan mag jij ook knallen tijdens je verdediging. Het was fijn en gezellig om je als kamergenoot te hebben.

Een speciaal woord van dank ook voor mijn (oud-) kamergenoten, Lise, Aniek, Nienke, Maartje en Maike. Bedankt voor de luisterende oren, het meedenken en de gezelligheid. Mede dankzij jullie reed ik met veel plezier naar Nijmegen.

Lieve Nienke, als oud-collega heb ik je leren kennen en wat was het leuk om een kamer met je te delen. Ik heb veel aan je gehad, we hebben veel met elkaar gedeeld en daarom ben ik ontzettend blij met de vriendschap die we er aan over hebben gehouden! Ik ben trots dat je me als paranimf wil helpen met de voorbereiding van mijn verdediging en dat we weer collega's zijn.

Mijn nieuwe collega's bedank ik voor het warme welkom. Ik heb veel zin om met jullie aan de slag te gaan.

Lieve volleybaldames, bedankt voor jullie interesse en vooral de nodige ontspanning tijdens trainingen, wedstrijden en gezellige uitjes!

Beste Lucky Seven dames, ook jullie bedankt voor de interesse en betrokkenheid de afgelopen jaren. De ervaringen uit de zorg die jullie met me delen zijn altijd een belangrijke inspiratiebron voor me geweest.

Lieve Nienke, Norie, Yvonne en Marjan (en natuurlijk ook jullie lieve mannen en kids), ik ken jullie al sinds de peuterspeelzaal en ben heel blij met jullie als vaste waarde in mijn leven. Het is ontzettend fijn dat we er altijd voor elkaar zijn als het mee zit, maar ook als het tegen zit. Dank voor jullie betrokkenheid en interesse in mijn werkzaamheden van de afgelopen jaren. Ik vind het fijn en belangrijk dat we, ondanks alle drukte van een baan en gezin, tijd vinden om elkaar regelmatig te zien, bij te kletsen over alle persoonlijke ontwikkelingen en leuke dingen te doen. Ik kijk met heel veel plezier terug op de wintersportweken, fietsweekenden, etentjes en de momenten met onze kids.

Ook wil ik graag mijn schoonfamilie, Bennie, Margriet, Sjors en Eva bedanken voor alle steun, interesse en gezelligheid de afgelopen jaren.

Lieve oma, Nijmegen vond je maar ver van huis. Afijn, als ik maar goed mijn best deed, dan was je blij. Bedankt voor je aanmoediging.

Pap en mam, ik had me geen beter thuis kunnen wensen. Bedankt dat jullie altijd klaar staan voor mij, Daan, Stan, aanhang en kleinkinderen. Het is bewonderenswaardig hoe jullie in het leven staan en klaar staan voor de mensen om jullie heen! Jullie vinden het belangrijk dat we doen wat we leuk vinden, het beste uit onszelf halen, niet te snel opgeven, nuchter blijven en vooral genieten van het leven. Bedankt voor alles wat jullie voor mij hebben gedaan!

Lieve Daan en Stan, jullie zijn me al lang boven het hoofd gegroeid maar toch blijven jullie mijn kleine broertjes! Ik ben heel blij met jullie en jullie lieve vriendinnen Luca en Remi. Sorry als ik af en toe nog steeds de grote zus uithang, maar ik ben hartstikke trots op wat jullie doen, hoe jullie in het leven staan en op jullie brede interesse en sociale persoonlijkheden. Zo, is dat ook eens gezegd! Het inspireert me en maakt me een blijer mens. Het is fijn te weten dat ik altijd op jullie kan rekenen. Dat ik binnenkort ook de titel 'tante' mag dragen maakt me extra trots.

Lieve Sjaak, jij bent degene die altijd in me gelooft, naar me luistert en mij het vertrouwen geeft om de beste versie van mezelf te zijn. Dank ook voor je engelengeduld, voor die ontelbare keren dat ik zei: "ik ben nog even 5 minuten wat af maken" terwijl jij ondertussen je lunchwandeling al lang had kunnen maken. Of alle avonden dat ik 's avonds weer aan het werk ging en niet te vergeten het bijkomend effect van de zwangerschapshormonen. Gelukkig voel jij me haarfijn aan en heb je me enorm geholpen om alle ballen in de lucht te houden. Onze vele herinneringen samen zijn me dierbaar. Ik kijk uit naar al het moois dat nog gaat komen, samen met onze meisjes, Roos en Merel, want jullie maken mij gelukkig!

Curriculum vitae

Curriculum vitae

Malou te Molder was born on May 8th, 1991 in Winterswijk and grew up in Zieuwent, the Netherlands. She started her academic journey after graduating from high school in 2008 (HAVO, Marianum, Groenlo). She completed her Nursing education at Saxion University of Applied Sciences in Enschede and obtained her bachelor's degree in 2012. Subsequently, she started a pre-master's degree in health sciences at the University of Twente, followed by a master's degree in health sciences with a specialization in epidemiology at Maastricht University. During her master's degree, Malou did an internship at the Epidemiology Department within the GROW School of Oncology and Developmental Biology at Maastricht University. Besides her studies, she worked as a nurse at Buurtzorg.

In 2015, Malou started as a research nurse in the orthopedics department of the Sint Maartenskliniek. After six months she combined her role as a research nurse with a position as a junior researcher.

In 2018, she expanded her activities at the Sint Maartenskliniek with her own PhD trajectory. She was supervised in this process by Dr. Petra Heesterbeek, Dr. Els van den Ende, Dr. José Smolders and Prof. Dr. Marinus de Kleuver. The different research projects aimed at the conceptualization of poor response after total knee arthroplasty are presented in this thesis.

Malou currently works as a quality policy officer in a healthcare institution. Malou lives in Zieuwent with her partner Sjaak and their daughters Roos (2021) and Merel (2023).

List of publications

List of publications

This thesis

Te Molder MEM, Smolders JMH, Heesterbeek PJC, Van den Ende CHM. Definitions of poor outcome after total knee arthroplasty: an inventory review. BMC Musculoskeletal Disord. 2020 Jun 13;21(1):378. doi: 10.1186/s12891-020-03406-y.

Te Molder MEM, Vriezekolk JE, Bénard MR, Heesterbeek PJC. Translation, cross-cultural adaptation, reliability and construct validity of the Dutch Oxford Knee Score – Activity and Participation Questionnaire. BMC Musculoskeletal Disord. 2021 Aug 17;22(1):700. doi: 10.1186/s12891-021-04521-0.

Te Molder MEM, Verhoef LM, Smolders JMH, Heesterbeek PJC, Van den Ende CHM. Prioritization of adverse consequences after total knee arthroplasty contributing to a poor response: A best-worst scaling exercise among total knee arthroplasty patients and knee specialists. J Arthroplasty. 2024 Mar 39(3):651. doi:10.1016/j.arth.2023.08.060.

Te Molder MEM, Vriezekolk JE, Van Onsem S, Smolders JMH, Heesterbeek PJC, Van den Ende CHM. Exploration of adverse consequences of total knee arthroplasty by patients and knee specialists: a qualitative study. Rheumatol Adv Pract. 2024;8(1):1-8. doi: 10.1093/rap/rkad111. eCollection 2024.

Te Molder MEM, Dowsey MM, Smolders JMH, Van Steenbergen LN, Van den Ende CHM, Heesterbeek PJC. Comparison of performance of different definitions of poor response after total knee arthroplasty using the Dutch Arthroplasty Register and the Osteoarthritis Initiative database. [Submitted]

Te Molder MEM, Van Onsem S, Smolders JMH, Dowsey MM, Rolfson O, Singh JA, De Kleuver M, Heesterbeek PJC, Van den Ende CHM. International consensus-based ranking of definitions for poor response to primary total knee arthroplasty: A Delphi study. [Submitted]

Other publications in peer-reviewed journals

Van Roekel EH, Bours MJL, **Te Molder MEM**, Breedveld-Peters JJL, Olde Damink SWM, Schouten LJ, Sanduleanu S, Beets GL, Weijenberg MP. Associations of adipose and muscle tissue parameters at colorectal cancer diagnosis with long-term health-related quality of life. Qual Life Res. 2017 Jul;26(7):1745. doi: 10.1007/s11136-017-1539-z.

Te Molder MEM, Wymenga AB, Heesterbeek PJC. Mid-flexion laxity in the asymptomatic native knee is predominantly present on the lateral side. Knee Surg Sports Traumatol Arthrosc. 2019 Nov; 27(11):3614. doi: 10.1007/s00167-019-05474-7.

Te Molder MEM, Smolders JMH, Heesterbeek PJC, Van den Ende CHM. Reply to letter to the editor: 'Definitions of poor outcome after total knee arthroplasty: an inventory review'. BMC Musculoskeletal Disord. 2020 Sep;21(1):604.

Van Laarhoven SN, **Te Molder MEM**, Van Hellemondt GG, Heesterbeek PJC. Acceptable migration of a fully cemented rotating hinge-type knee revision system measured in 20 patients with model-based RSA with a 2-year follow-up. Acta Orthop. 2023 Apr 24;94:185. doi: 10.2340/17453674.2023.12305.

Invited presentations

Te Molder MEM. Ontevreden na een nieuwe knie. Congres Verder in beweging Nijmegen, 2022.

Te Molder MEM. Ontevreden na een knieprothese. NERASS Najaarscongres Woerden, 2022.

Conference abstracts thesis-related

Te Molder MEM, Smolders JMH, Heesterbeek PJC, Van den Ende CHM. Identification of definitions of poor outcome after treatment of knee osteoarthritis: A literature review. EULAR Madrid, digital, 2019 (abstract book)

Te Molder MEM, Smolders JMH, Heesterbeek PJC, Van den Ende CHM. How to define poor response to primary total knee replacement? EKS Arthroplasty Conference Valencia, 2019 (poster presentation); EULAR congress Madrid, 2019 (poster)

Te Molder MEM, Vriezekolk JE, Van Onsem S, Smolders JMH, Heesterbeek PJC, Van den Ende CHM. Poor response to primary total knee arthroplasty: the perspective of patients and knee specialists. EULAR congress Virtual Congress, 2021 (poster presentation); EKS Arthroplasty Conference Munich, 2022 (oral presentation)

Te Molder MEM, Dowsey MM, Smolders JMH, Van Steenbergen LN, Heesterbeek PJC, Van den Ende CHM. Comparison of prevalence and overlap of different definitions of poor response after total knee arthroplasty using the Dutch Arthroplasty Register and the Osteoarthritis Initiative database. ISAR Annual Congress Montreal, 2023 (oral presentation); EKS Arthroplasty Conference Munich, 2022 (poster presentation)

Te Molder MEM, Van Onsem S, Smolders JMH, Dowsey MM, Rolfson O, Singh JA, De Kleuver M, Heesterbeek PJC, Van den Ende CHM. International consensus-based criteria for poor response to primary total knee arthroplasty: A Delphi study. EKS Arthroplasty Conference Munich, 2022 (oral presentation)

Te Molder MEM, Verhoef LM, Smolders JMH, Heesterbeek PJC, Van den Ende CHM. Prioritization of consequences after total knee arthroplasty contributing to a poor response: A best-worst scaling exercise among total knee arthroplasty patients and knee specialists. ISTA Annual Congress New York, 2023 (oral presentation); NOV Najaarscongres Rotterdam, 2023 (oral presentation); NVR Najaarsdagen Arnhem, 2023 (pitch)

Other conference abstracts

Te Molder MEM, Wymenga AB, Heesterbeek PJC. Normal knee laxity at 0º, 30º, and 90º of flexion measured in the older healthy knee. NOV Najaarscongres Veldhoven, 2017 (oral presentation); EKS Arthroplasty Conference London (poster)

Te Molder MEM, Heesterbeek PJC, Wymenga AB, Van Hellemondt GG. Micromotion of a cemented hinged type knee revision system with model-based RSA. 2nd World Arthroplasty Congress Rome, 2018 (poster presentation)

Research Data Management

Research Data Management

All studies in this thesis were presented to the medical and ethical review board Committee on Research Involving Human Subjects Region Arnhem-Nijmegen, the Netherlands. For all studies the review board provided a waiver as all studies fell outside the remit of the law for Medical Research Involving Human Subjects Act and were approved by the local ethical committees. Additionally, the Belgium part of the qualitative study in Chapter 4, was approved by the institutional ethics committee of the Ghent University Hospital (BC-07096). In addition, the study in Chapter 5 did not require approval from an ethical committee in the Netherlands according to the Medical Research Involving Human Subject Act since the researchers had access only to unidentifiable patient data in both the LROI register and the OAI database. All data were handled in line with the principles of the declaration of Helsinki. Informed consent was obtained from all study participants prior to any study procedures when applicable.

The privacy of the participants in all studies is warranted by use of encrypted and unique individual subject codes. The code was stored separately from the study data. Data were converged to Stata (StataCorp LLC, College Station, Texas, USA) and SPSS (IBM, Armonk, New York, USA) for analyses. Data from the qualitative study in Chapter 4 was analyzed using Atlas.ti (Scientific Software Development GmbH). Data from the best-worst scaling exercise in Chapter 7 was collected online by the use of Sawtooth Software.

The data are digitally stored on the research servers (H:\) and (V:\) of the Sint Maartenskliniek and on paper in the research department's archive. The data will be saved for 20 years after termination of the studies and are available through the corresponding authors upon reasonable request.

PhD Portfolio

PhD portfolio of M.E.M. te Molder

Department: Research, Sint Maartenskliniek
PhD period: 01/01/2018 – 30/06/2023
PhD Supervisor(s): Prof. dr. M. de Kleuver, Dr. C.H.M. van den Ende
PhD Co-supervisor(s): Dr. P.J.C. Heesterbeek, Dr. J.M.H. Smolders

Training activities	Hours
Courses	
- Introduction to model-based RSA course (2016) LUMC RSAcore	8.00
- BROK course (2016) NFU BROK Academie	26.00
- Workshop Discrete Choice Experiments (2017)	8.00
- Scientific Writing for PhD candidates (2018)	84.00
- Advanced Conversation (2018)	42.00
- Kwalitatief onderzoek in de praktijk van de gezondheidszorg (K78) (2018) EpidM	56.00
- RIHS - Introduction course for PhD candidates (2019)	15.00
- Workshop statistiek (2019)	4.00
- Radboudumc - Scientific integrity (2019)	20.00
- Science Journalism and Communication (2020)	84.00
- Advanced model-based RSA course (2020) LUMC RSAcore	8.00
- Re-registration BROK course (2020)	5.00
- Klinimetrie: het ontwikkelen en evalueren van meetinstrumenten (V40) (2020) EpidM	54.00
- Design and Illustration (2020)	28.00
- 'How to write a peer-review?' (2021)	4.00
- 'Structure and focus in your text' (2021)	2.00
- 'Strategies for revising your own text' (2021)	2.00
- 'Writing a rebuttal' (2021)	4.00
- Radboudumc - General Radboudumc introduction for research personnel (2023)	9.00
Seminars	
- Masterclass patiënt participatie (2017)	8.00
- Scientific integrity lecture (Lex Bouter) (2018)	6.00
- Work visit from George Peat (Keele University) (oral presentation) (2019)	4.00
Conferences	
- EKS Arthroplasty Conference (2017)	18.00
- Third Joint Meeting BKS/DKS (2017)	8.00
- NOV conference (oral presentation) (2017)	21.00
- 2nd World Arthroplasty Congress (oral presentation) (2018)	30.00
- EKS Arthroplasty Conference (oral presentation) (2019)	21.00
- EULAR congress (poster presentation) (2019)	30.00
- Verder in beweging conference Sint Maartenskliniek (2016-2020)	40.00
- EULAR congress (oral presentation) (2021)	24.00
- PhD retreat RIHS (2021)	16.00
- EKS Arthroplasty Conference (two oral presentations and poster presentation) (2022)	24.00
- Verder in beweging conference Sint Maartenskliniek (invited oral presentation) (2022)	24.00
- Nerass conference (invited oral presentation) (2022)	20.00
- ISAR conference (oral presentation) (2023)	6.00
- NVR conference (poster presentation) (2023)	6.00
- ISTA conference (oral presentation) (2023)	6.00
- NOV conference (oral presentation) (2023)	6.00
Other	
- Journal club Sint Maartenskliniek (2018-2023)	35.00
- Research lunch Sint Maartenskliniek (2018-2023)	40.00

- Peer-review of scientific publications (2019-2023)	14.00
Teaching activities	
Supervision of internships / other	
- Supervision research project 6 nursing students (2018)	28.00
Total	**898.00**

Theses Sint Maartenskliniek

Theses Sint Maartenskliniek

Kuijpers, R. (2024). *Adapt your step: Clinical assessment and training of walking adaptability in children with mild motor disorders.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Den Broeder, A. (2024). *More than tapering, less than full dose - Efficient use of biologics in the treatment of rheumatoid arthritis.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Veenstra, F. (2024). *About gout. Studying potential targets for improvement of care.* Radboud University Nijmegen, Nijmegen. The Netherlands.

De Jong, L.A.F. (2023). *Effects of lower limb orthotic devices in people with neurological disorders.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Michielsens, C. (2023). *Tapering strategies of biologics in inflammatory disorders*. Radboud University Nijmegen, Nijmegen. The Netherlands.

Pouls, B. (2023). *Supporting patients'medication management using eHealth. Test cases in rheumatology*. Radboud University Nijmegen, Nijmegen. The Netherlands.

Stöcker, J. (2023). *Accessible and effective non-pharmacological care for persons with systemic sclerosis*. Radboud University Nijmegen, Nijmegen. The Netherlands.

Mulder, M. (2022). *Going off-road. Exploring and mapping psoriatic arthritis.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Huiskes, V. (2022). *The synergistic role of patients and healthcare providers in reducing drug-related problems.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Marsman, D. (2022). *Polymyalgia rheumatica. Clinical characteristics and new treatment opportunities.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Alingh, J. (2021). *Effect of robotic gait training on the post-stroke gait pattern. Evaluation of LOPES II.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Van Dijsseldonk, R. (2021). *Step into the future: mobility after spinal cord injury*. Radboud University Nijmegen, Nijmegen, The Netherlands.

Pelle, T. (2021). *Beating osteoarthritis by e-self management in knee or hip osteoarthritis.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Van Heuckelum, M (2020). *Novel approaches to improve medication adherence in rheumatoid arthritis.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Mathijssen, E. (2020). *The voice of patients with rheumatoid arthritis.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Bakker, S. (2019). *Regional anesthesia and total knee arthroplasty. Anesthetic and pharmacological considerations.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Claassen, A. (2019). *Strategies for patient education in rheumatic diseases.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Fenten, M. (2019). *Optimizing locoregional anesthesia in fast track orthopaedic surgery.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Minten, M. (2019). *On the role of inflammation and the value of low dose radiation therapy in osteoarthritis.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Verhoef, L. (2019). *Effective and efficient use of bDMARDs in rheumatoid arthritis.* Radboud University Nijmegen, Nijmegen. The Netherlands.

Bekker, C. (2018). *Sustainable use of medication. Medication waste and feasibility of redispensing.* Utrecht University, Utrecht. The Netherlands.

Bikker, I.(2018). *Organizing timely treatment in multi-disciplinary care.* University of Twente, The Netherlands.

Bouman, C. (2018). *Dose optimisation of biologic DMARDs in rheumatoid arthritis: long-term effects and possible predictors.* Radboud University Nijmegen, The Netherlands.

Mahler, E. (2018). *Contributors to the management of osteoarthritis.* Utrecht University, The Netherlands.

Tweehuysen, L. (2018). *Optimising biological treatment in inflammatory rheumatic diseases. Predicting, tapering and transitioning.* Radboud University Nijmegen, Nijmegen, The Netherlands.

Geerdink, Y. (2017). *Getting a grip on hand use in unilateral cerebral palsy.* Radboud University, Nijmegen, The Netherlands.

Remijn, L. (2017). *Mastication in children with cerebral palsy.* Radboud University, Nijmegen, The Netherlands.

Selten, E. (2017). *Beliefs underlying treatment choices in osteoarthritis.* Radboud University, Nijmegen, The Netherlands.

Van Hooff, M. (2017). *Towards a paradigm shift in chronic low back pain? Identification of patient profiles to guide treatment.* VU University Amsterdam, Amsterdam, The Netherlands.

Lesuis, N. (2016). *Quality of care in rheumatology. Translating evidence into practice.* Radboud University, Nijmegen, The Netherlands.

Luites, J. (2016). *Innovations in femoral tunnel positioning for anatomical ACL reconstruction.* Radboud University, Nijmegen, The Netherlands.

Pakvis, D. (2016). *Survival, primary stability and bone remodeling assessment of cementless sockets. An appraisal of Wolff's law in the acetabulum*. Radboud University, Nijmegen, The Netherlands.

Schoenmakers, K. (2016). *Prolongation of regional anesthesia. Determinants of peripheral nerve block duration.* Radboud University, Nijmegen, The Netherlands.

Altmann, V. (2015). *Impact of trunk impairment on activity limitation with a focus on wheelchair rugby.* Leuven University, Leuven, Belgium.

Bevers, K. (2015). *Pathophysiologic and prognostic value of ultrasonography in knee osteoarthritis.* Utrecht University, Utrecht, The Netherlands.

Cuperus, N. (2015). *Strategies to improve non-pharmacological care in generalized osteoarthritis.* Radboud University, Nijmegen, The Netherlands.

Kilkens, A. (2015). *De ontwikkeling en evaluatie van het Communicatie Assessment & Interventie Systeem (CAIS) voor het aanleren van (proto-)imperatief gedrag aan kinderen met complexe ontwikkelingsproblemen.* Radboud University, Nijmegen, The Netherlands.

Penning, L. (2015). *The effectiveness of injections in cuffdisorders and improvement of diagnostics.* Maastricht University, Maastricht, The Netherlands.

Stegeman, M. (2015). *Fusion of the tarsal joints: outcome, diagnostics and management of patient expectations.* Utrecht University, Utrecht, The Netherlands.

Van Herwaarden, N. (2015). *Individualised biological treatment in rheumatoid arthritis.* Utrecht University, Utrecht, The Netherlands.

Wiegant, K. (2015). *Uitstel kunstknie door kniedistractie.* Utrecht University, Utrecht, The Netherlands.

Willems, L. (2015). *Non-pharmacological care for patients with systemic sclerosis.* Radboud University, Nijmegen, The Netherlands.

Witteveen, A. (2015). *The conservative treatment of ankle osteoarthritis.* University of Amsterdam, Amsterdam, The Netherlands.

Zwikker, H. (2015). *All about beliefs. Exploring and intervening on beliefs about medication to improve adherence in patients with rheumatoid arthritis.* Radboud University, Nijmegen, The Netherlands.

Koenraadt, K. (2014). *Shedding light on cortical control of movement.* Radboud University, Nijmegen, The Netherlands.

Smink, A. (2014). *Beating Osteoarthritis. Implementation of a stepped care strategy to manage hip or knee osteoarthritis in clinical practice.* VU University Amsterdam, Amsterdam, The Netherlands.

Stolwijk, N. (2014). *Feet 4 feet. Plantar pressure and kinematics of the healthy and painful foot.* Radboud University, Nijmegen, The Netherlands.

Van Kessel, M. (2014). *Nothing left? How to keep on the right track. Spatial and non-spatial attention processes in neglect after stroke.* Radboud University, Nijmegen, The Netherlands.

Brinkman, M. (2013). *Fixation stability and new surgical concepts of osteotomies around the knee.* Utrecht University, Utrecht, The Netherlands.

Kwakkenbos, L. (2013). *Psychological well-being in systemic sclerosis: Moving forward in assessment and treatment.* Radboud University, Nijmegen, The Netherlands.

Severens, M. (2013). *Towards clinical BCI applications: assistive technology and gait rehabilitation.* Radboud University, Nijmegen, The Netherlands.

Stukstette, M. (2013). *Understanding and treating hand osteoarthritis: a challenge.* Utrecht University, Utrecht, The Netherlands.

Van der Maas, A. (2013). *Dose reduction of TNF blockers in Rheumatoid Arthritis: clinical and pharmacological aspects.* Radboud University, Nijmegen, The Netherlands.

Zedlitz, A. (2013). *Brittle brain power. Post-stroke fatigue, explorations into assessment and treatment.* Radboud University, Nijmegen, The Netherlands.

Beijer, L. (2012). *E-learning based speech therapy (EST). Exploring the potentials of E-health for dysarthric speakers.* Radboud University, Nijmegen, The Netherlands.

Hoogeboom, T. (2012). *Tailoring conservative care in osteoarthritis.* Maastricht University, Maastricht, The Netherlands.

Boelen, D. (2011). *Order out of chaos? Assessment and treatment of executive disorders in brain-injured patients.* Radboud University, Nijmegen, The Netherlands.

Heesterbeek, P. (2011). *Mind the gaps! Clinical and technical aspects of PCL-retaining total knee replacement with the balanced gap technique.* Radboud University, Nijmegen, The Netherlands.

Hegeman, J. (2011). *Fall risk and medication. New methods for the assessment of risk factors in commonly used medicines.* Radboud University, Nijmegen, The Netherlands.

Smulders, E. (2011). *Falls in rheumatic diseases. Risk factors and preventive strategies in osteoporosis and rheumatoid arthritis.* Radboud University, Nijmegen, The Netherlands.

Snijders, G. (2011). *Improving conservative treatment of knee and hip osteoarthritis.* Radboud University, Nijmegen, The Netherlands.

Vriezekolk, J. (2011). *Targeting distress in rheumatic diseases.* Utrecht University, Utrecht, The Netherlands.

Willems, P. (2011). *Decision making in surgical treatment of chronic low back pain. The performance of prognostic tests to select patients for lumbar spinal fusion.* Maastricht University, Maastricht, The Netherlands.

Aarts, P. (2010). *Modified constraint-induced movement therapy for children with unilateral spastic cerebral palsy: the Pirate group intervention.* Radboud University, Nijmegen, The Netherlands.

Groen, B. (2010). *Martial arts techniques to reduce fall severity.* Radboud University, Nijmegen, The Netherlands.

Van Koulil, S. (2010). *Tailored cognitive behavioral therapy in fibromyalgia.* Radboud University, Nijmegen, The Netherlands.

Van den Bemt, B. (2009). *Optimizing pharmacotherapy in patients with rheumatoid arthritis: an individualized approach.* Radboud University, Nijmegen, The Netherlands.

Van Nes, I. (2009). *Balance recovery after supratentorial stroke. Influence of hemineglect and the effects of somatosensory stimulation.* Radboud University, Nijmegen, The Netherlands.

Ruiter, M. (2008). *Speaking in ellipses. The effect of a compensatory style of speech on functional communication in chronic agrammatism.* Radboud University, Nijmegen, The Netherlands.

Baken, B. (2007). *Reflexion on reflexes. Modulation during gait.* Radboud University, Nijmegen, The Netherlands.

Gaasbeek, R. (2007). *High tibial osteotomy. Treatment of varus osteoarthritis of the knee.* Radboud University, Nijmegen, The Netherlands.

Koëter, S. (2007). *Patellar instability. Diagnosis and treatment.* Radboud University, Nijmegen, The Netherlands.

Langeloo, D. (2007). *Monitoring the spinal cord during corrective spinal surgery: a clinical study of TES-MEP.* Radboud University, Nijmegen, The Netherlands.

De Haart, M. (2005). *Recovery of standing balance in patients with a supratentorial stroke.* Radboud University, Nijmegen, The Netherlands.

Den Otter, R. (2005). *The control of gait after stroke: an electromyographic approach to functional recovery.* Groningen University, Groningen, The Netherlands.

Spruit, M. (2005). *Surgical treatment of degenerative disc conditions of the lumbar spine. Biomechanical, clinical and radiological aspects.* University Utrecht, Utrecht, The Netherlands.

Weerdesteyn, V. (2005). *From the mechanisms of obstacle avoidance towards the prevention of falls.* Radboud University, Nijmegen, The Netherlands.

Jongerius, P. (2004). *Botulinum toxin type-A to treat drooling. A study in children with cerebral palsy.* Radboud University, Nijmegen, The Netherlands.

Van de Crommert, H. (2004). *Sensory control of gait and its relation to locomotion after a spinal cord injury.* Radboud University, Nijmegen, The Netherlands.

Van der Linde, H. (2004). *Prosthetic prescription in lower limb amputation. Development of a clinical guideline in the Netherlands.* Groningen University, Groningen, The Netherlands.

Hendricks, H. (2003). *Motor evoked potentials in predicting motor and functional outcome after stroke.* University of Nijmegen, Nijmegen, The Netherlands.

Hosman, A. J. F. (2003). *Idiopathic thoracic spinal deformities and compensatory mechanisms.* University of Nijmegen, Nijmegen, The Netherlands.

Donker, S. (2002). F*lexibility of human walking: a study on interlimb coordination.* Groningen University, Groningen, The Netherlands.

Hochstenbach, J. (1999). *The cognitive, emotional, and behavioural consequences of stroke.* University of Nijmegen, The Netherlands.

De Kleuver, M. (1998). *Triple osteotomy of the pelvis. An anatomical, biomechanical and clinical study.* University of Nijmegen, Nijmegen, The Netherlands.

Van Balen, H. (1997). *A disability-oriented approach to long-term sequelae following traumatic brain injury. Neuropsychological assessment for post-acute rehabilitation*. University of Nijmegen, Nijmegen, The Netherlands.

Tromp, E. (1995). *Neglect in action: a neuropsychological exploration of some behavioural aspects of neglect*. University of Nijmegen, Nijmegen, The Netherlands.

Van Lankveld, W. (1993). *Coping with chronic stressors of rheumatoid arthritis.* University of Nijmegen, Nijmegen, The Netherlands.

Geurts, A. (1992). *Central adaptation of postural organization to peripheral sensorimotor impairments.* University of Nijmegen, Nijmegen, The Netherlands.

De Rooij, D. (1988). *Clinical and serological studies in the connective tissue diseases.* University of Nijmegen, Nijmegen, The Netherlands.

Sint Maartenskliniek

Radboudumc
university medical center

Radboud University