EXPECTATIONS AND OUTCOME IN KNEE & HIP ARTHROPLASTY

Jaap J. Tolk

EXPECTATIONS AND OUTCOME IN KNEE & HIP ARTHROPLASTY

Jaap J. Tolk

Jaap Tolk
Expectations and Outcome in Knee and Hip Arthroplasty
Verwachtingen en uitkomst bij knie- en heupprothesen

ISBN 978-94-6332-625-4

Cover & layout by Beatrijs Voorneman

Printed by GVO drukkers & vormgevers B.V.

Financial support by the Nederlandse Orthopedische Vereniging (NOV) is gratefully acknowledged. Part of the research described in this thesis was supported by a grant from the van Rens Funds.

EXPECTATIONS AND OUTCOME IN KNEE AND HIP ARTHROPLASTY

Verwachtingen en uitkomst bij knie- en heupprothesen

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. R.C.M.E. Engels

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

Dinsdag 2 Juni 2020 om 13.30 uur

door

Jaap Johannes Tolk
geboren te Venlo

Erasmus University Rotterdam

Promotiecommissie:

Promotor:	Prof.dr. S.M.A. Bierma-Zeinstra
Overige leden:	Prof.dr. L.J.W. van Rhijn Prof.dr. J.M.W. Hazes Prof.dr. B.W. Koes
Copromotoren:	Dr. M. Reijman Dr. R.P.A. Janssen

Stellingen

behorende bij het proefschrift 'Expectations and Outcome in Knee and Hip Arthroplasty'

1. De functietesten van de OARSI-core set zijn niet geschikt voor het meten van fysiek functioneren in de klinische praktijk. (dit proefschrift)

2. Patiënten kunnen gemiddeld twee weken na een totale heupprothese en vier weken na een totale knieprothese weer autorijden. (dit proefschrift)

3. Vrouwelijk geslacht, hogere leeftijd, hogere depressie scores en langere klachtenduur zijn geassocieerd met lagere verwachtingen van het behandelresultaat na een knieprothese. (dit proefschrift)

4. Patiënten die aanvullend preoperatief verwachtingsmanagement krijgen hebben meer uitgekomen verwachtingen en een hogere postoperatieve tevredenheid na een totale knieprothese. (dit proefschrift)

5. Demografische gegevens en PROMs die verzameld worden voor de LROI, kunnen gebruikt worden voor een geïndividualiseerde voorspelling van het behandelresultaat. (dit proefschrift)

6. Big data gaat een grote rol spelen bij individuele klinische beslissingen, het blijft echter de taak van de dokter om de inherente onzekerheid van een voorspelling te duiden. (Chen, 2017)

7. De patiënt en zijn persoonlijke omstandigheden hebben meer invloed op tevredenheid dan de operateur. (Khanna, 2019)

8. Ondanks het stijgende aantal fietsongevallen, leef je van fietsen toch langer. (de Hartog, 2010)

9. Studenten moeten op weg naar het einddiploma in de gelegenheid zijn om voldoende studiepunten én levenspunten te behalen. (Nieuwenhuijzen Kruseman, 2010)

10. Crises bring out the best in people – 'andrà tutto bene'. (Quarantelli 2008)

11. Stilte is het verschil tussen niks zeggen en alles al gezegd hebben. (Herman de Coninck)

Jaap J. Tolk
Rotterdam, 2 Juni 2020

table of **Contents**

chapter 1

General introduction

Osteoarthritis (OA) is the most common musculoskeletal disease.[1] In 2018 1,467,200 people were estimated to have OA in the Netherlands; 513,900 men and 953,300 women.[1] Knee OA was most prevalent, with approximately 693,400 patients affected.[1]

OA is primarily characterised by articular cartilage loss, but the whole joint is affected with possible changes to subchondral bone, osteophyte formation, bone deformation and synovial membrane reaction.[2] The main symptoms are joint pain and stiffness, resulting in restrictions in activities of daily living and a negative influence on quality of life.[3] Knee and hip OA can result in considerable physical and psychological impairment and is a debilitating condition in end-stages of the disease.[4]

Main treatment goals in the treatment of OA are pain relief and improvement of physical function.[5] Primarily the treatment consists of non-operative measures such as activity level adjustment, exercise therapy and medication.[6] When conservative treatment has proven to be insufficient effect, joint replacement can be considered.[7] But, especially for knee OA several options are available, including osteotomies, knee joint distraction, unicompartmental or total knee arthroplasty.[8,9] The choice between treatment modalities should be made depending on patient characteristics, anatomical parameters, affected compartments and patient preferences.[7,8]

Total knee arthroplasty (TKA) is the most frequently performed procedure for patients with knee OA. The Dutch arthroplasty registry reported an increase from 18.500 TKA in 2010, to 25.269 in 2018 in the Netherlands.[10] The same trend is observed for the number of THA procedures in the Netherlands. This has increased from 23.340 THA in 2010, to 31.599 in 2018.[10] A further increase in these numbers is expected in the future due to aging of the Western population and growing number of people with obesity.[11]

Outcome assessment after joint arthroplasty

Total knee and hip arthroplasty are generally considered successful treatments for patients with knee or hip OA, respectively, but 'treatment success' is a multi-interpretable term. Traditionally outcome parameters such as prosthesis alignment, survival, postoperative range of motion and complication percentage are most frequently reported. Outcome in this regard is generally successful; the treatment is relatively safe, cost-effective and excellent survival rates are reported, with prosthesis survival of more than 95% at 15 years follow-up.[12–14]

In addition to surgeon oriented and implant specific outcome parameters, nowadays patient reported outcome measures (PROMs) and patient satisfaction are increasingly implemented as criteria of treatment success.[15–17] Overall these self-reported outcomes can be considered good as well; considerable pain reduction, increase in physical function and quality of life can be achieved.[13,18] On the other hand, especially in measurement of physical function concerns have been raised with respect to the limitation of solely relying on PROMs to assess this domain.[19–22]

In OA research it is advocated to evaluate the effect of a treatment on change in pain, function and patient's global assessment.[23] Whereas for pain and global assessment PROMs are generally considered the method of choice, for the measurement of physical function more options are available.[16] Both self-reported measures of function and instruments that directly asses the execution of a specific task associated with function (performance-based tests) are available.[24] PROMs aim to assess a patients perception of their physical functioning, whereas performance-based measures aim to quantify the performance of a specific activity.[25] A discrepancy in results after TKA between these two methods is reported[19,26], leading to the idea that different aspects of the construct physical functioning are measured.[19–21] Furthermore, performance-based measures would be less pain-driven and suffer less from ceiling effects, when compared to self-reported measures of physical function.[19,26] Complementary application in the evaluation of physical function in an assessment continuum is suggested.[16,22,24] Nevertheless, the available evidence on the measurement properties of available performance-based measures is limited for knee and hip OA patients. This warrants further investigation before broad clinical application should be considered.[21]

Expectations and satisfaction

Despite the generally favourable results concerning pain reduction and improvement in physical function, the rate of satisfaction after TKA is consistently reported around 80%.[15,27] This leaves approximately 1 in 5 patients unsatisfied to some extent after their knee surgery. One of the main determinants of post-operative satisfaction is reported to be the fulfilment of pre-operative outcome expectations.[15,27–30]

Patients planned for TKA have multiple expectations regarding the most likely treatment result. Most expectations concern relief of pain, improvement in physical functioning and improvement in psychosocial well-being.[31,32] A discrepancy often exists between expectations of the patients and those of the surgeon.[33] Surgeons generally have lower expectations of the most likely treatment result.[33] Patients tend to have high expectations and, as a substantial number of patients is reported to have unfulfilled expectations after TKA, they often seem to be too optimistic.[30,34,35]

In a study by Hamilton et al. analysing factors affecting postoperative satisfaction after primary TKA in 2247 patients, the main predictor of satisfaction was 'meeting preoperative expectations' [29]. Bourne et al. found similar results in a study of 1703 primary TKA, where patients with expectations that were not met, were at 10.7x greater risk to be dissatisfied with treatment outcome [28]. Fulfilment of expectations is reported as the strongest predictor of treatment satisfaction, with more influence than pain relief, postoperative complications and pre- or postoperative physical status [28,29]. These findings support the expectancy-disconfirmation theory, which states that satisfaction is a function of expectations, perceived performance, and disconfirmation of beliefs.[36]

Expectation management

Considering the strong relationship between expectation fulfilment and satisfaction, expectation management in TKA patients aimed at realistic postoperative expectations, is thought advantageous to achieve optimal patient satisfaction.

In this thesis we mainly refer to 'probabilistic outcome expectations'. Patient expectations are defined as 'anticipations that given events are likely to occur during or as a result of medical care', [37] and expectations on the result of a treatment as 'outcome expectations'.[38] For these outcome expectation two dimensions can be distinguished: value-based and probabilistic outcome expectations. Value-based outcome expectations concern what a patient considers to be most important, and are thought to be mainly emotionally driven, reflecting desire, hopes and wishes. Probalistic outcome expectations on the other hand, address what a patient thinks will be the most likely result of treatment and are more cognitively driven.[39] These constructs can be distinguished in TKA patients and are often not aligned.[40] In the light of expectation management interventions, ideally a patients' value-based outcome expectation should determine the specific subjects that are emphasized in the pre-operative education, subsequently for these items a patients probability-based expectations should be aligned with the most likely treatment result.

Considering value-based expectations, previous research has identified a set of expectations that are considered most important by TKA patients.[31,32] The most important items concern pain relief and improvement in functioning in daily life (e.g. walking, chair rising, stair climbing), performing social activities (e.g. hobbies, sport activities) and psychological well-being. Although these factors are considered important by most patients, there seems to be considerable variance over population and age groups.[41] Concerning probability based outcome expectations there seems to be large individual differences as well. Age, and sex are reported as significant independent predictors of expectations, and it is suggested that psychological factors and personality traits may play significant roles in outcome expectations.[42,43] Still, the available evidence on patient factors that determine expectations of patients awaiting TKA is limited. More insight in factors that determine patients' expectations can be useful to guide pre-operative education and the decision-making process.

Previously it has been shown that pre-operative education addressing realistic expectations for long term recovery can change patients' pre-operative expectations.[44] Pre-operative education is reported to result in lower patient expectations, and a higher concordance in patients' and surgeons' expectations.[33,44] These findings suggest a beneficiary effect of enhanced pre-operative expectation management, but the effect on post-operative expectation fulfilment and ultimately better post-operative satisfaction after TKA has not yet been confirmed.

Currently available structured education modules mainly describe realistic expectations for the general population of patients undergoing arthroplasty.[33,44] When an individualised outcome prediction would be possible, this could be very useful in pre-operative education.

It has been shown that useful prediction on postoperative outcome can be made from pre-operative patient factors.[45–48] Existing outcome prediction tools for knee OA mainly focus on identifying patients most likely not to benefit from TKA.[46,49,50] Specific information on pain and functional outcome to guide pre-operative expectation management is not provided by these tools.[46,49,50] For effective expectation management, a prediction tool should ideally provide specific information on pain and functional outcome for an individual patient. This would make prediction tools a valuable asset in improved pre-operative education on realistic expectations for TKA patients.

Overview of the content of this thesis

Part 01 | Outcome assessment after arthroplasty

The first part of this thesis focuses on measurement of physical function after joint arthroplasty. The Osteoarthritis Research Society International (OARSI) recommends a set of performance-based tests to assess the construct physical function, based on expert opinion.[16,21] Nevertheless, evidence on the measurement properties of the performance-based measures included in this set is limited.[21,25]

In **Chapter 2 and 3** we assessed the measurement properties of the OARSI recommended performance-based measures for measurement of physical function in patients with severe knee OA (chapter 2) and severe hip OA (chapter 3).

Part 02 | Expectations of treatment result

The second part of this thesis addresses what can be considered realistic expectations for treatment result after TKA and analysed determinants of patients' expectations.

Chapter 4 presents a survey among Dutch orthopaedic surgeons, addressing what these experts assume are realistic expectations for long-term recovery after total knee arthroplasty.

Patients consider return to driving independently after knee or hip arthroplasty as an important factor in postoperative recovery, as it increases mobility and reduces social isolation and dependence on others. Nevertheless, there is no consensus on when it is safe to drive after THA or TKA. In a systematic review in **Chapter 5** we aimed to assess the current available evidence about when patients might resume driving after elective, primary THA or TKA

It is recognized that fulfilment of expectations plays an important role in determining postoperative satisfaction. [15,27–30] There is limited evidence on what determines the level of patients' expectations. Psychological factors and personality traits may play significant roles, in addition to demographic factors, pain, physical function and general health.[32,43,51] The aim of the study presented in **Chapter 6** was to analyse the relationship between pre-operative factors and pre-operative outcome expectations in TKA patients.

Part 03 | Expectation management in clinical practice

Pre-operative expectation management to improve postoperative patient satisfaction has not yet been translated into a successful intervention. In the third part of this thesis we aimed to come to clinically applicable modalities, harnessing the potential of improved expectation management for TKA patients.

In **Chapter 7** elaborates on the working mechanism behind an expectation modification intervention and the protocol for a randomised controlled trial (RCT) assessing the effect of an additional expectation management module for TKA patients is presented.

Chapter 8 presents the results of this RCT examining whether an additional education module on realistic expectations for long-term recovery of symptoms, physical functioning and psychological issues (intervention group) would improve patient satisfaction after TKA compared to usual pre-operative education (control group).
In **Chapter 9** the development and validation of prediction models for outcome after TKA are presented, based on data gathered for registration in the Dutch Arthroplasty register. The models aim to predict the chance of residual symptoms after TKA for an individual patient on 10 specific items concerning treatment success, functional outcome and pain relief.

Finally, in **Chapter 10** a general discussion is presented on the main findings, future research perspectives and implications for clinical practice of the studies described in this thesis.

References

1. Prevalentie en aantal nieuwe gevallen van artrose. https://www.volksgezondheidenzorg.info/onderwerp/artrose/cijfers-context/huidige-situatie. Accessed November 14, 2019.
2. Loeser F, Goldring S, Scanzello C, Goldring M. Osteoarthritis: A Disease of the Joint as an Organ Richard. *Arthritis Rheum*. 2012;64(6):1697-1707. doi:10.1038/jid.2014.371
3. Hoy D, Cross M, Smith E, et al. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis*. 2014;73(7):1323-1330. doi:10.1136/annrheumdis-2013-204763
4. Litwic A, Edwards MH, Dennison EM, Cooper C. Epidemiology and burden of osteoarthritis. *Br Med Bull*. 2013;105(1):185-199. doi:10.1093/bmb/lds038
5. 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. *Osteoarthr Cartil*. 2008;16(2):137-162. doi:10.1016/j.joca.2007.12.013
6. Bannuru RR, Osani MC, Vaysbrot EE, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. *Osteoarthr Cartil*. 2019;epub ahead. doi:10.1016/J.JOCA.2019.06.011
7. Nederlandse Orthopaedische Vereniging. Richtlijn totale knieprothese. 2014.
8. McGrory BJ, Weber KL, Jevsevar DS, Sevarino K. Surgical management of osteoarthritis of the knee: Evidence-based guideline. *J Am Acad Orthop Surg*. 2016;24(8):e87-e93. doi:10.5435/JAAOS-D-16-00159
9. van der Woude J, Wiegant K, van Heerwaarden R, et al. Knee joint distraction compared with total knee arthroplasty: a randomised controlled trial. *Bone Jt J*. 2017;99-B(1):51-58.
10. Dutch Arthroplasty Register (LROI). *LROI Annual Report 2019*. 's Hertogenbosch, The Netherlands; 2019.
11. Dutch Arthroplasty Register (LROI). *Online LROI Annual Report 2017: 10 Years of Registration, a Wealth of Information*. 's-Hertogenbosch, The Netherlands; 2017.
12. McCalden RW, Hart GP, MacDonald SJ, Naudie DD, Howard JH, Bourne RB. Clinical Results and Survivorship of the GENESIS II Total Knee Arthroplasty at a Minimum of 15 Years. *J Arthroplasty*. 2017;32(7):2161-2166. doi:10.1016/j.arth.2017.02.006
13. Ruiz D, Koenig L, Dall TM, et al. The Direct and Indirect Costs to Society of Treatment for End-Stage Knee Osteoarthritis. *J Bone Jt Surg Am*. 2013;95(16):1473-1480. doi:10.2106/JBJS.L.01488
14. Räsänen P, Paavolainen P, Sintonen H, et al. Effectiveness of hip or knee replacement surgery in terms of quality-adjusted life years and costs. *Acta Orthop*. 2007;78(1):108-115. doi:10.1080/17453670610013501
15. 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):148-152. doi:10.1302/0301-620x.95b11.32767
16. Dobson F, Hinman RS, Roos EM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. *Osteoarthr Cartil*. 2013;21(8):1042-1052. doi:10.1016/j.joca.2013.05.002
17. Tolk J, Koot H, Janssen R. Computer Navigated Versus Conventional Total Knee Arthroplasty. *J Knee Surg*. 2012;25(04):347-352. doi:10.1055/s-0031-1299670
18. Tolk JJ, Janssen RPA, Prinsen CAC, et al. The OARSI core set of performance-based measures for knee osteoarthritis is reliable but not valid and responsive. *Knee Surgery, Sport Traumatol Arthrosc*. 2019;27(9):2898-2909. doi:10.1007/s00167-017-4789-y
19. Mizner RL, Petterson SC, Clements KE, Zeni J a, Irrgang JJ, Snyder-Mackler L. Measuring

functional improvement after total knee arthroplasty requires both performance-based and patient-report assessments: a longitudinal analysis of outcomes. *J Arthroplasty*. 2011;26(5):728-737. doi:10.1016/j.arth.2010.06.004

20. Stratford PW, Kennedy DM, Woodhouse LJ. Performance measures provide assessments of pain and function in people with advanced osteoarthritis of the hip or knee. *Phys Ther*. 2006;86(11):1489-1496. doi:10.2522/ptj.20060002
21. Dobson F, Hinman RS, Hall M, Terwee CB, Roos EM, Bennell KL. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis : a systematic review. *Osteoarthr Cartil*. 2012;20(12):1548-1562. doi:10.1016/j.joca.2012.08.015
22. Stratford PW, Kennedy DM. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. *J Clin Epidemiol*. 2006;59(2):160-167. doi:10.1016/j.jclinepi.2005.07.012
23. 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):389-399. doi:10.1016/j.joca.2004.02.001
24. Reiman MP, Manske RC. The assessment of function: How is it measured? A clinical perspective. *J Man Manip Ther*. 2011;19(2):91-99. doi:10.1179/10669811 1X12973307659546
25. Terwee CB, Mokkink LB, Steultjens MPM, Dekker J. Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Rheumatology (Oxford)*. 2006;45(7):890-902. doi:10.1093/rheumatology/kei267
26. Stevens-Lapsley JE, Schenkman ML, Dayton MR. Comparison of self-reported knee injury and osteoarthritis outcome score to performance measures in patients after total knee arthroplasty. *PM R*. 2011;3(6):541-549; quiz 549. doi:10.1016/j.pmrj.2011.03.002
27. Noble PC, Conditt M a, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452(452):35-43. doi:10.1097/01.blo.0000238825.63648.1e
28. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: Who is satisfied and who is not? *Clin Orthop Relat Res*. 2010;468(1):57-63. doi:10.1007/s11999-009-1119-9
29. Hamilton DF, Lane J V, Gaston P, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open*. 2013;3(4):1-8. doi:10.1136/bmjopen-2012-002525
30. 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 Orthop*. 2009;80(1):55-61. doi:10.1080/17453670902805007
31. Mancuso CA, Sculco TP, Wickiewicz TL, Jones EC, Robbins L, Warren RF. Patients' expectations of knee surgery. *J Bone Jt Surg Am*. 2001;83-A(7):1005-1012.
32. Smith EJ, Soon V-L, Boyd A, McAllister J, Deakin A, Sarungi M. What do Scottish Patients Expect of Their Total Knee Arthroplasty? *J Arthroplasty*. 2016;31(4):786-792. doi:10.1016/j.arth.2015.10.021
33. Ghomrawi HMK, Mancuso C a., Westrich GH, Marx RG, Mushlin AI. Discordance in TKA expectations between patients and surgeons knee. *Clin Orthop Relat Res*. 2013;471(1):175-180. doi:10.1007/s11999-012-2484-3
34. Mannion AF, Kämpfen S, Munzinger U, Kramers-de Quervain I. The role of patient expectations in predicting outcome after total knee arthroplasty. *Arthritis Res Ther*. 2009;11(5):R139. doi:10.1186/ar2811

35. Tilbury C, Haanstra TM, Leichtenberg CS, et al. Unfulfilled Expectations After Total Hip and Knee Arthroplasty Surgery: There Is a Need for Better Preoperative Patient Information and Education. *J Arthroplasty*. 2016;31(10):2139-2145. doi:10.1016/j.arth.2016.02.061
36. Oliver RL. Effect of Expectation and Disconfirmation on Postexposure Product Evaluations: An Alternative Interpretation. *J Appl Psychol*. 1977;62(4):480-486.
37. Uhlmann RF, Inui TS, Carter WB. Patient requests and expectations. Definitions and clinical applications. *Med Care*. 1984;22(7):681-685. doi:10.1097/00005650-198407000-00011
38. Haanstra TM, van den Berg T, Ostelo RW, et al. Systematic review: do patient expectations influence treatment outcomes in total knee and total hip arthroplasty? *Health Qual Life Outcomes*. 2012;10(1):152. doi:10.1186/1477-7525-10-152
39. Kravitz RL. Patients' expectations for medical care: an expanded formulation based on review of the literature. *Med Care Res Rev*. 1996;53(1):3-27.
40. Haanstra TM, Tilbury C, Kamper SJ, et al. Can Optimism, Pessimism, Hope, Treatment Credibility and Treatment Expectancy Be Distinguished in Patients Undergoing Total Hip and Total Knee Arthroplasty? *PLoS One*. 2015;10(7):e0133730. doi:10.1371/journal.pone.0133730
41. Yoo JH, Kang YG, Kim SJ, Seong SC. Patient expectations of total knee replacement and their association with sociodemographic factors and functional status. *J Bone Joint Surg Br*. 2011;93(3):337-344. doi:10.1302/0301-620X.93B3.25168
42. Razmjou H, Finkelstein JA, Yee A, Holtby R, Vidmar M, Ford M. Relationship between preoperative patient characteristics and expectations in candidates for total knee arthroplasty. *Physiother Canada*. 2009;61(1):38-45. doi:10.3138/physio.61.1.38
43. Hepinstall MS, Rutledge JR, Bornstein LJ, Mazumdar M, Westrich GH. Factors that impact expectations before total knee arthroplasty. *J Arthroplasty*. 2011;26(6):870-876. doi:10.1016/j.arth.2010.09.010
44. Mancuso C, Graziano S, Briskie L, et al. Randomized trials to modify patients' preoperative expectations of hip and knee arthroplasties. *Clin Orthop Relat Res*. 2008;466(2):424-431. doi:10.1007/s11999-007-0052-z
45. Dowsey MM, Spelman T, Choong PFM. Development of a Prognostic Nomogram for Predicting the Probability of Nonresponse to Total Knee Arthroplasty 1 Year After Surgery. *J Arthroplasty*. 2015;31:1654-1660. doi:10.1016/j.arth.2016.02.003
46. 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. doi:10.1016/j.arth.2016.11.007
47. Pua YH, Seah FJT, Clark RA, Poon CLL, Tan JWM, Chong HC. Development of a prediction model to estimate the risk of walking limitations in patients with total knee arthroplasty. *J Rheumatol*. 2016;43(2):419-426. doi:10.3899/jrheum.150724
48. 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;31:1-9. doi:10.1016/j.arth.2016.06.004
49. Barlow T, Dunbar M, Sprowson A, Parsons N, Griffin D. Development of an outcome prediction tool for patients considering a total knee replacement--the Knee Outcome Prediction Study (KOPS). *BMC Musculoskelet Disord*. 2014;15(1):451. doi:10.1186/1471-2474-15-451
50. Lungu E, Desmeules F, Dionne CE, Belzile EL, Vendittoli P-A. Prediction of poor outcomes six months following total knee arthroplasty in patients awaiting surgery. *BMC Musculoskelet Disord*. 2014;15:299. doi:10.1186/1471-2474-15-299
51. Jain D, Nguyen L-CL, Bendich I, et al. Higher patient expectations predict higher patient-reported outcomes, but not satisfaction, in total knee arthroplasty patients: a prospective

Outcome assessment after arthroplasty

part 01

Abstract

Purpose

The Osteoarthritis Research Society International (OARSI) has identified a core set of performance-based tests of physical function for use in people with knee osteoarthritis (OA). The core set consists of the 30-second chair-stand test (30s CST), 4x10 meter fast-paced walk test (40m FPWT) and a stair climb test. The aim of this study was to evaluate the reliability, validity and responsiveness of these performance-based measures to assess the ability to measure physical function in knee OA patients.

Methods

A prospective cohort study of 85 knee OA patients indicated for total knee arthroplasty (TKA) was performed. Construct validity and responsiveness were assessed by testing of predefined hypotheses. A subgroup (n=30) underwent test-retest measurements for reliability analysis. The Oxford Knee Score, Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form, pain during activity score and knee extensor strength were used as comparator instruments. Measurements were obtained at baseline and 12 months after TKA.

Results

Appropriate test-retest reliability was found for all three tests. Intraclass Correlation Coefficient (ICC) for the 30s CST was 0.90 (95% CI 0.68;0.96), 40m FPWT 0.93 (0.85;0.96) and for the 10 step Stair Climb Test (10-step SCT) 0.94 (0.89;0.97). Adequate construct validity could not be confirmed for the three tests. For the 30s CST 42 % of the predefined hypotheses were confirmed, for the 40m FPWT 27 % and for the 10-step SCT 36 % confirmed. The 40m FPWT was found to be responsive with 75% of predefined hypothesis confirmed whereas the responsiveness for the other tests could not be confirmed. For the 30s CST and 10-step SCT only 50% of hypotheses were confirmed.

Conclusions

The three performance-based tests had good reliability, but poor construct validity and responsiveness in the assessment of function for the domains sit-to-stand movement, walking short distances and stair negotiation. The findings of the present study do not justify their use for clinical practice.

chapter 2

The OARSI core set of performance-based measures for knee osteoarthritis is reliable but not valid and responsive.

J.J. Tolk, R.P.A. Janssen, C.A.C. Prinsen, D.A.J.M. Latijnhouwers,
M.C. van der Steen, S.M.A. Bierma-Zeinstra, M. Reijman

Knee Surgery, Sports Traumatology and Arthroscopy. 2019 Sep;27(9):2898-2909.

Introduction

Knee osteoarthritis (OA) has large societal, psychological and physical burdens for patients affected by the disease.[1] Knee OA patients often experience pain and restrictions in physical functioning.[2] Important goals of knee OA treatment with total knee arthroplasty (TKA) are pain relief and improvement of physical function.[3]

The evaluation of treatment outcome after TKA should at least assess the domains pain, function and a global assessment.[4] For the measurement of physical function, self-reported measures of function and testing of the execution of a specific task associated with function (performance-based tests) can be used.[5] Whereas patient reported outcome measures (PROMs) assess what patients perceive they can do, performance-based measures aim to quantify what patients actually can do.[6] When measuring change in physical function after TKA, a discrepancy is observed between the results of these methods.[7,8] This leads to the idea that these two types of outcome measurement instruments, although being related, measure different aspects of the construct physical functioning.[7,9,10] Integration of both types of measurement in an assessment continuum is suggested, and considered complementary in the evaluation of physical function. [5,11]

The functional tasks that are most relevant to measure are pathology- and population specific.[12] The three most relevant functional domains for knee OA are level walking, stair negotiation and sit-to-stand movement.[13] Impairment on these domains is classified as 'activity limitations' on the World Health Organisation International Classification of Functioning, Disability and Health (ICF).[14]

Based on currently available evidence and expert consensus the Osteoarthritis Research Society International (OARSI) identified a set of performance-based tests to assess these functional domains.[10,13] The aimed construct of measurement is physical function, which is related to the ability to "move around" and "perform daily activities" and can be classified as Activities using the ICF model.[10,13,14] The core set consist of the 30-s chair-stand test (30s CST), 4x10 meter fast-paced walk test (40m FPWT) and a stair-climb test.[13]

For tests to be usable in both clinical practice and research, measurement properties should be appropriate.[15,16] Data on the reliability, validity and responsiveness of the OARSI core set of performance-based measures is either unavailable or from low quality studies.[10] Therefore good quality research investigating measurement properties of these performance measures is necessary.[6,10] The aim of the current study was to evaluate the reliability, validity and responsiveness of the core set performance-based measures for measurement of physical function in knee OA patients.

Materials and methods

A prospective cohort study of patients indicated for TKA was performed. Evaluation of measurement properties of the 30s CST, 40m FPWT and 10-step chair climb test (10-step SCT) was conducted following the COSMIN methodology (COnsensus based Standards

for the selection of health status Measurement INstruments).[16] The COSMIN checklist is a consensus-based checklist and can be used to evaluate the methodological quality of studies on measurement properties of health status measurement instruments.[16] The Máxima MC Medical Ethics Committee approved the study (registration code 2014-73).

Patient population

All symptomatic knee OA patients scheduled for primary TKA in Máxima MC were eligible for inclusion. Exclusion criteria were comorbidity leading to inability to perform the performance-based tests, insufficient knowledge of the Dutch language leading to inability to fill out the study questionnaires and inability to visit follow-up appointments. If the patient met the criteria and was willing to participate, an informed consent form was signed.

Study procedures

At baseline the following clinical parameters were recorded; side of operation, gender, age, and body mass index (BMI).

Testing procedures took place at the outpatient clinic of Máxima MC, in a designated testing area by a research nurse. Measurement of the OARSI core set of performance-based tests was executed strictly according to the manual provided by the OARSI, following a standardized protocol with the following fixed order of tests.[13] Measurements were obtained pre-operatively and 12 months postoperative.

Performance-based measures

30s CST

The 30s CST is a performance-based measure that evaluates the activity 'sit-to-stand movement'.[13] The test is executed by scoring the maximum amount of complete chair stand movements during 30 seconds. A full sit-to-stand and consecutive stand-to-sit cycle is counted as one chair stand. A 43 cm high, straight back chair without arm rests was used. To date no previous reliability reports specifically for knee OA patients are available. In a combined group of hip and knee OA patients excellent reliability is reported, with an Intraclass Correlation Coefficients (ICC) of 0.95 (SD 0.93-0.97), and a Standard Error of Measurement (SEM) of 0.7 repetitions.[17] Construct validity and responsiveness have not been reported previously in knee OA patients.

40m FPWT

The 40m FPWT assesses the activity 'walking short distances'.[13] It scores the maximal walking speed on a marked walkway of 4 times 10 meters, excluding turns. The result is expressed as speed in meters / second (m/s). There are no previous reports on the reliability of this version of the 40m FPWT.[18] Kennedy et al. report on a similar walk test, scoring walking speed using a walkway of 2 times 20-meter. Their results show good reliability with an ICC of 0.91 (SD 0.81, 0.97) and SEM of 1.73 m/s (SD 1.39-2.29).[18] No previous reports on construct validity of the 40m FPWT are available in literature.[10]

Stair Climb Test

For assessment of the activity 'stair negotiation', no specific stair climb test is advised by the OARSI.[13] In the present study, the 10-step stair climb test (10-step SCT) was selected, as the stair in the testing area had 10 steps. The step height was 18.8 cm and depth 22.4 cm. The

time needed to ascend and descent these steps is recorded in seconds. No previous reports on reliability of the 10-step SCT are available. Almeida et al. reported excellent reliability with an ICC of 0.94 (SD 0.55-0.98) and a SEM of 2.35s for the 11-step stair test in knee OA patients.[19] The 11-step SCT is essentially the same test as the 10-step version, with the only difference that the stairway used has one step more.

Comparator instruments

KOOS-PS

The Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form (*KOOS-PS*) Dutch version is a 7-item questionnaire that assesses the construct physical function. From a 5-point Likert scale question, a normalized score is calculated (0 indicating no symptoms and 100 indicating extreme symptoms).[20] KOOS-PS has good reliability, face and content validity and ability to detect change over time in knee OA patients.[20–23]

OKS

The Dutch version of the Oxford Knee Score (OKS) is a 12-item PROM designed to measure function and pain after TKA. Each question consists of a 5-point Likert scale, leading to a total score ranging from a best functional score of 12 to the worst functional outcome of 60.[24] It is short, reproducible, valid and sensitive to clinically important changes.[24] The OKS has adequate internal consistency and test retest reliability, good face, content and construct validity and good sensitivity and responsiveness in knee OA patients.[23]

EQ-5D

The Dutch version of the EuroQol 5D-3L (EQ-5D) is a 5-item PROM, measuring generic health status.[25] Scoring the lowest score on the EQ-5D index indicates the worst health state possible and a score of 1 represents the best possible health state.[25] The EQ-5D has good reliability and validity in knee OA patients.[26]

NRS pain

Numerical Rating Scale (NRS) for pain during activity (NRS pain) was used to measure level of pain during activity. The scale consists of eleven points in which the patient can score the pain during activities in general from 0 to 10. A score of 0 represented 'no pain' and a score of 10 represented 'worst imaginable pain'. The NRS has good reliability and responsiveness.[21]

Anchor question

At 12 months postoperative follow-up a 7-point Likert scale anchor question was scored for change in activities of daily living. Response options ranged from 1 (a lot worse) to 7 (very much improved).

ROM

Range of motion (ROM) of the affected knee was measured in supine position using a goniometer, considering the bony landmarks of the greater trochanter, lateral femoral condyle, and lateral malleolus. Maximal flexion was scored as positive value and an extension deficit was scored as negative value. In knee OA, ROM measurement has adequate reliability with a reliability coefficient of 0.81 for extension and 0.96 for flexion.[27]

Quadriceps strength

To determine Quadriceps strength of the affected leg, maximal isometric knee extensor strength was measured using a handheld dynamometer (HHD). Testing took place in an upright position. The HHD was positioned perpendicular to the anterior aspect of the tibia, 5 cm proximal of the medial malleolus. A protective shin guard was used for patient comfort as well as standardisation of HHD placement. 3 consecutive measurements were obtained, of which the highest value was used for analysis. An HHD is a widely used, reliable, and valid instrument to measure knee extensor strength, with good reliability in OA patients (ICC 0.94).[28]

Evaluation of the measurement properties

Reliability

Reliability is defined as the extent to which scores for patients, who have not changed, are the same for repeated measurement under similar conditions.[16] To evaluate the reliability of the 3 performance-based tests, test-retest measurements were obtained in a random subset of patients. After initial measurement (To) patients rested for 30 minutes, after which a second round of testing was performed (To_1). This test-retest design was considered appropriate as the resting period allows full recovery from the performed tests, and the tested function can be assumed to remain stable over the testing period. Circumstances, setting, order of the 3 tests and instructions in the retest setting were identical to the first round of testing. Reliability analysis consisted of determining ICC for absolute agreement with corresponding 95% Confidence Intervals (CI), SEM, and Smallest Detectable Change (SDC). An ICC value > 0.70 is considered appropriate.[29,30]

Construct validity

There is no 'Gold Standard' available for assessment of the functional domains level walking, stair negotiation and sit-to-stand movement in knee OA. Therefore, determining construct validity is the designated method to analyse the degree to which the studied measurement instruments are measuring the constructs that they aim to measure.15,16,31 This method is internationally accepted and recommended by the COSMIN for these circumstances.15,16,31 Predefined hypotheses were formulated on the relationships of performance-based tests scores with scores on other instruments measuring similar or dissimilar constructs.[15,31] A panel comprising of four experts in the field of outcome measurement in knee OA (orthopedic surgeon, orthopedic resident and Ph.D. candidate, specialist in measurement property analysis and methodologist), formulated 11 to 15 hypotheses for each measurement instrument under study. An overview of the hypotheses can be found in Table 3.

The predefined hypotheses consisted of both convergent and discriminant validity hypotheses, and comparative hypothesis on a closer relationship with similar compared to dissimilar constructs. The hypothesis included direction and magnitude of the expected results. In general, we hypothesized the following. The performance-based measures would be moderately correlated to PROMs and quadriceps strength. PROMs have a stronger correlation with pain scores than with the performance-based measures. Performance-based measures were expected to have a stronger correlation with PROMs measuring functional outcome than with a PROM measuring general health. Specific questions of the PROMs regarding walking, stair negotiation and sit-to-stand movement were expected to correlate stronger to their respective performance-based measure than to the total score of the PROM.

Correlations of measurements with similar constructs were expected to be at least moderate ≥ 0.4 or ≤ -0.4. Measurements that were unrelated or had different constructs were expected to have a poor correlation [-≥0.39; ≤0.39]. The performance-based tests are assumed valid if at least 75% of the predefined hypotheses are confirmed.[29,30]

Responsiveness

Responsiveness is defined as the ability of the instruments to detect change over time in the construct measured.[15,16,31] In the absence of a gold standard, the assessment of responsiveness relies on hypotheses testing (i.e. a construct approach).[15,16,31] These hypotheses concern the expected relationships between changes on the studied instruments and changes on other instruments that measure similar or different constructs with adequate responsiveness.[15,29,31] These hypothesis, with expected direction and magnitude of the correlations, were formulated a priori.

The performance-based tests are assumed to be adequately responsive if minimally 75% of the predefined hypotheses are confirmed.[29,30] The responsiveness hypothesis can be found in Table 5. In summary, it was hypothesised that the anchor question was moderately correlated to change in the performance-based measures scores. Only a moderate correlation was expected, because experienced change in functional ability is not exactly the same construct as actual change in execution of the task. Furthermore, we hypothesised that the change in PROMs is more correlated to pain, than to change in the performance-based test scores.

Statistical analysis

Statistical analysis was performed with SPSS statistics version 24.0 (IBM corporation). The reliability analysis was performed using a Two-Way Random Model with absolute agreement. SEM was calculated using the formula: Standard Deviation (SD) difference / √n. Where n represents the number of measurement repetitions; n=2 for the present study. The SDC was calculated as 1.96 × √2 × SEM.[32] For the construct validity and responsiveness analysis Pearson or Spearman correlation coefficients were calculated, depending on normality of data distribution. Comparison of Performance-based measures and PROM scores before and after TKA was conducted using a paired samples t-test or Wilcoxon signed-rank test, depending on normality of data distribution. The sample size was based on the COSMIN criteria, aiming for a good score for the construct validity and responsiveness analysis (≥ 50 patients) and fair for reliability assessment (≥ 30 patients).[16,29]

Table 1. Baseline characteristics

	Total cohort (n=85)		Reliability analysis cohort (n=30)	
Age, years	69.3	(± 8.2)	67.8	(±7.7)
Gender, female n (%)	46	(57)	13	(43)
Side affected, right n (%)	41	(48)	17	(57)
BMI, kg/m2	29.6	(± 5.0)	29.9	(±5.6)
Maximal flexion, degrees	110	(± 17.0)	106	(±18.9)
Extension deficit, degrees	4	(± 7.0)	4	(±6.5)

Data are presented as mean and standard deviation between parentheses, or reported otherwise as mentioned.

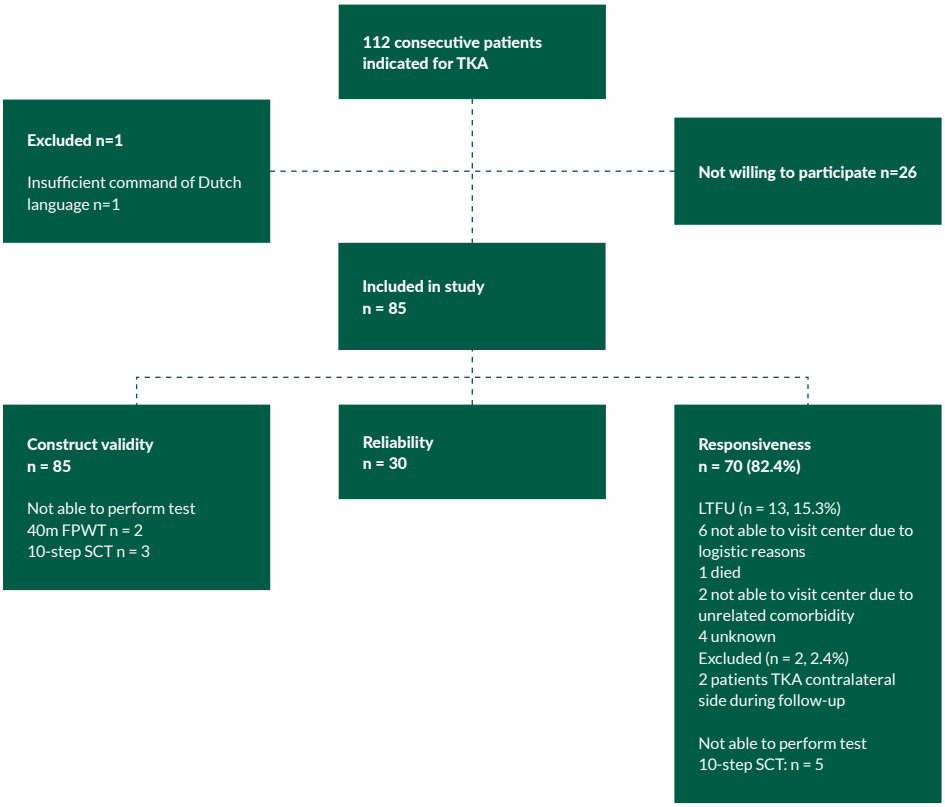


Figure 1. *Number of patients included in analysis and reasons for loss to follow up. LTFU, Lost to follow up.*

Results

Patient characteristics

Between April and October 2015, 85 consecutive patients with knee OA were included. The baseline characteristics are described in Table 1. Number of patients included in the reliability, construct validity and responsiveness analysis and reasons for loss to follow-up are summarised in Figure 1.

Measurement properties

a) Reliability analysis

Test-retest measurements were performed in a random subgroup consisting of the first 30 patients that were included in the study. Mean test scores and reliability parameters are presented in Table 2.

b) Construct validity (hypothesis testing)

Spearmans' correlation coefficients for the construct validity analysis are presented in Table 3. Confirmation of 75% or more of the predefined hypotheses was achieved by none of the three performance-based measures. 5/12 (42%) were confirmed for the 30s CST, 4/15 (27%) for the 40m FPWT and 4/11 (36%) for the 30s SCT.

Table 2. Reliability analysis (n = 30)

	Mean score baseline	Mean retest score	Mean of Difference (baseline - retest score)	ICC	SEM	SDC
30s CST (stands)	9.0 (7.9-10.1)	9.8 (8.6-11.0)	0.8 (0.4-1.3)	0.90 (0.68-0.96)	0.85	2.4
40m FPWT (m/s)	1.30 (1.16-1.44)	1.32 (1.20-1.45)	- 0.02 (-0.08-0.03)	0.93 (0.85-0.96)	0.10	0.27
10-step SCT (seconds)	16.7 (13.5-19.9)	16.6 (13.5-19.7)	0.1 (-1.0-1.1)	0.94 (0.89-0.97)	1.98	5.5

ICC, Intraclass Correlation Coefficient; SEM, Standard Error of Measurement; SDC, Smallest Detectable Change. Data are presented as mean and 95% Confidence Interval between parentheses, or reported otherwise as mentioned.

Table 4. Performance-based measures and PROM scores before and after TKA

	Baseline	12 month follow-up after TKA	p-value
30s CST (stands)	9.2 (8.4-10.0)	11.3 (10.3-12.4)	<0.001
40m FPWT (m/s)	1.25 (1.16-1.34)	1.38 (1.25-1.50)	0.001
Use of assistive device during 40m FPWT (patients, n)	2	0	NA
10-step SCT (seconds)	21.8 (18.4-25.1)	15.5 (13.9-17.1)	0.007
Use of handrail 10-step SCT (patients, n)	39	24	0.40 (n.s.)
KOOS-PS score	54.2 (50.8-57.5)	28.9 (24.6-33.1)	<0.001
OKS	21.7 (20.2-23.2)	40.1 (38.1-42.1)	<0.001
EQ5D	0.48 (0.42-0.55)	0.84 (0.79-0.89	<0.001
NRS pain	7.6 (7.2-7.9)	2.1 (1.6-2.7)	<0.001

Data are presented as mean and 95% Confidence Interval between parentheses, or reported otherwise as mentioned. n.s., non-significant.

c) Responsiveness

The scores of the performance-based measures at baseline and after TKA at 12-month follow up are presented in Table 4. All performance-based measures, PROMs and the NRS pain score showed significant improvement at 12-month follow-up. Only the use of a handrail during the 10-step SCT did not show significant change. On the anchor question for change in activities of daily living the mean score at 12 month follow up was 6.2 (95%CI 5.9-6.5), this represents 'much improved'. Spearmans' correlation coefficients for responsiveness analysis are presented in Table 5. For the 30s CST 4/8 (50%) of the hypothesis were confirmed, for the 40m FPWT 6/8 (75%) and for the 10-step SCT 4/8 (50%).

Discussion

The present study showed good reliability of the OARSI recommended core set of performance-based measures. However, based on a low percentage of confirmation of our predefined hypotheses, construct validity and responsiveness of the tests was poor.

Test-retest reliability of the three performance-based measures is adequate, as the presented ICC values are well above 0.70, which is considered acceptable.[33] This is in line with previous reports on test-retest reliability for these tests.[18,19] The SDC values reported in the present study for the 30s CST and 10-step SCT are similar to those reported in literature.[17,19] There is no consensus on what SDC value is acceptable.[32] From a clinical point of view, the SDC's of 2.5 stands for the CST and 0.27 m/s for the 40m FPWT reported in the present study seem reasonable. This is different however for the 10-step SCT. With an SDC of 5.5 seconds, an individual patient has to improve or deteriorate almost 1/3 of the mean time taken for the initial test, to be certain a change has occurred. From a clinical perspective, this seems quite a large difference, resulting in a low sensitivity to change on the tested functional domain.

In the construct validity assessment, the necessary 75% hypotheses confirmation was achieved by none of the performance-based tests. The main reason for this was the rejection of all convergent hypotheses for correlations between the performance-based measures and the patient reported measures of function. As PROMs are, by definition, subjective measures of function, only a moderate correlation with the more objective measurements of the performance-based measures was expected. For example, PROMs are known to be more related to pain than to actual execution of the task at hand,[7,8,34] as was also found in the present study. Self-reported and performance based assessment of activities are inherently linked, considering that both methods aim to measure the same 'activities' defined in the ICF theoretical framework.[14] In our view for performance-based measures to be clinically relevant, some relation between experienced performance and the result of the performance-based measure of this activity should be present. However, even the moderate correlations we predicted were not met, resulting in poor construct validity.

An explanation for the poor construct validity might be that timed measures of performance did not fully capture impairment on the activities at hand. The time taken to execute a task is not the only factor in the performance of this task in daily living. A patient might execute the activity swiftly, but if the quality of performance is affected by for example limping or

instability, it can still be considerably impaired.[11,35,36] Such an impairment cannot be captured by merely timing the activity.[35,36] Another explanation for discordance between self-reported and performance-based measurement of function can be underrepresentation.[37] Whereas the OKS and KOOS-ps measure the general construct physical function, the performance-based tests under study aim to quantify performance on specific functional tasks. The narrower construct of the performance test might not be represented by these two PROMs used as comparative instruments.[37] If underrepresentation were the case, the specific questions addressing the functional tasks measured by the respective tests would be likely to correlate stronger to these tests. To account for this, we made hypothesis on correlations with these specific questions. The correlations found on these hypotheses were even lower, making underrepresentation as an explanation unlikely.

The strong relationship between pain and self-reported function found in the construct validity analysis was even more obvious in the responsiveness analysis. The change in NRS pain score was strongly related to the change in subjective scores, but unrelated to the performance-based measures. This supports claims that performance-based measures are less pain driven than PROMs, and provide a more objective view on the task performed.[7,8] On the other hand, it is our opinion that for a test to be clinically relevant some relationship between actual change and experienced change in performance on the functional task at hand should exist. Therefore, we hypothesised that the overall change in PROM scores one year after TKA would correlate moderately to the change in performance-based measures. Only for the 40m FPWT most hypothesis in this regard were confirmed. For the other two tests, no such relationship was found. As mentioned earlier for the construct validity, underrepresentation and the inability of timed measures to fully capture impairment on the tested domains might explain the lack of responsiveness of the 30s CST and 10-step SCT.

A remark has to be made on the comparative instruments used for the construct validity and responsiveness analysis. These consisted of a combination of objective and subjective measurements of function and general health with good reliability, construct validity and responsiveness in a knee OA population.[38–41] Other options for comparison could have been objective measures such as optoelectric or inertia based motion analysis systems. These measures are suitable for a strictly kinematic analysis, but their clinical relevance has not been clarified.[35,42] Therefore, we believe that they are not suitable of the construct validity analysis in this regard. In our view, the comparative measurement instruments in the present study were the most appropriate instruments available.

To our knowledge, this is the first study assessing the most important measurement properties of the OARSI recommended core set of performance-based measures. A strength of the present study is the strictly followed, state-of-the-art methodology.[16] We report on an unselected, consecutive group of knee OA patients awaiting total knee arthroplasty in a general hospital. Previous reports on measurement properties often included a combined population of knee and hip OA patients, resulting in a more heterogeneous population.[17,18] Combining these distinctly different groups reduces the accuracy of the previously reported data. Our findings can be considered representative for knee OA patients.

The sample size can be considered good for the construct validity and responsiveness analysis and fair for reliability assessment.[16,29] A limitation of this research was the incomplete 12-month follow-up, the 82.4% follow-up achieved is however acceptable. For the subset of patients with incomplete data, no difference in preoperative demographics or baseline measurement was observed. Therefore, a systematic bias because of the loss to follow up seems unlikely. The results for the reliability assessment should be interpreted with some caution as a subgroup of only 30 patients was used. There is concurrent evidence on test-retest measurements from others studies, with similar results.[18,19] When combining these data, stronger evidence for an adequate reliability can be obtained. As mentioned earlier, the SDCs in the present study are relatively large, especially for the 10-step SCT. Test-retest measurements in a larger population would have resulted in a more precise determination of the SDC; it might be smaller than reported in the present study.

Conclusion

The OARSI core set of performance-based measures was advised to obtain a more complete view of the functional performance of knee OA patients.[13] The 30s CST, 40m FPWT and 10-step SCT have good reliability, but poor construct validity and responsiveness in the assessment of function and change in function for the domains sit-to-stand movement, walking short distances and stair negotiation respectively. The findings of the present study do not justify their use for clinical practice.

Acknowledgements

We would like to sincerely thank Christa van Doesburg, Hilke Cox and Mathias Mariam for their work in administrative and testing procedures.

Table 3. Construct validity.						
	30s CST		40m FPWT		10-step SCT (CC in opposite direction*)	
Predefined hypotheses	**Correlation Coefficient**	**Hypothesis Confirmed**	**Correlation Coefficient**	**Hypothesis Confirmed**	**Correlation Coefficient***	**Hypothesis Confirmed**
1. Moderate correlation with KOOS-PS (≤ -0.4)*	-0.33	No	-0.25	No	0.26	No
2. Moderate correlation with OKS (≥ 0.4)*	0.35	No	0.32	No	-0.33	No
3. Moderate correlation with Quadriceps strength (≥ 0.4)*	0.60	Yes	0.64	Yes	-0.74	Yes
4. Unrelated with EQ-5D [-0.35; 0.35]	0.25	Yes	0.18	Yes	-0.18	Yes
5. Correlation with KOOS-PS is minimal 0.1 stronger than with EQ-5D	-0.33 / 0.25	No	-0.25 / 0.18	No	0.26 / -0.18	No
6. Correlation with OKS is minimal 0.1 stronger than with EQ-5D	0.35 / 0.25	Yes	0.32 / 0.18	Yes	-0.33 / -0.18	Yes
7. 'Absolute' correlation between NRS pain and KOOS-PS is minimal 0.1 higher than between performance-based measure and NRS pain	0.37 / -0.10	Yes	0.37 / -0.07	Yes	0.37 / 0.01	Yes
8. 'Absolute' correlation between NRS pain and OKS is minimal 0.1 higher than performance-based measure and NRS pain	-0.45 / -0.10	Yes	NA		NA	
9. 'Absolute' correlation 30s CST with KOOS-PS Question 3 is minimal 0.1 higher than with KOOS-PS (total score)	-0.21 / -0.33	No	NA		NA	
10. 'Absolute' correlation 30s CST with KOOS-PS Question 3 is minimal 0.1 higher than with OKS	-0.21 / 0.35	No	NA		NA	
11. 'Absolute' correlation 30s CST with KOOS-PS Question 3 is minimal 0.1 higher than with EQ-5D Score	-0.21 / 0.25	No	NA		NA	
12. Moderate correlation 30s CST with KOOS-PS Question 3 (≤ -0.4)	-0.21	No	NA		NA	
13. 'Absolute' correlation 40m FPWT with EQ-5D Question 1 is minimal 0.1 stronger than with KOOS-PS	NA		-0.09 / 0.26	No	NA	

14. 'Absolute' 40m FPWT with EQ-5D Question 1 is minimal 0.1 stronger than with Oxford Knee Score	NA		-0.09 / 0.32	No	NA	
15. 'Absolute' correlation 40m FPWT with EQ-5D Question 1 is minimal 0.1 higher than with EQ-5D Score	NA		-0.09 / 0.18	No	NA	
16. 'Absolute' correlation 40m FPWT with OKS Question 6 is minimal 0.1 stronger than with KOOS-PS	NA		-0.03 / -0.25	No	NA	
17. 'Absolute' correlation 40m FPWT with OKS Question 6 is minimal 0.1 stronger than with OKS	NA		-0.03 / 0.32	No	NA	
18. 'Absolute' correlation 40m FPWT with OKS Question 6 is minimal 0.1 stronger than with EQ-5D Score	NA		-0.03 / 0.18	No	NA	
19. Moderate correlation 40m FPWT with EQ-5D Question 1 (≤ -0.4)	NA		-.09	No	NA	
20. Moderate correlation 40m FPWT with OKS Question 6 (≤ -0.4)	NA		-0.03	No	NA	
21. 'Absolute' correlation 10-step SCT with OKS Question 12 is minimal 0.1 stronger than with KOOS-PS	NA		NA		0.22 / 0.26	No
22. 'Absolute' correlation 10-step SCT with OKS Question 12 is minimal 0.1 stronger than with OKS	NA		NA		0.22 / -0.33	No
23. 'Absolute' correlation 10-step SCT with OKS Question 12 is minimal 0.1 stronger than with EQ-5D	NA		NA		0.22 / -0.18	No
24. Moderate correlation 10-step SCT with OKS Question 12 (≤-0.4)	NA		NA		0.22	No
Hypothesis confirmed	**5/12**	**42%**	**4 / 15**	**27%**	**4 / 11**	**36%**

*KOOS-PS; Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form, OKS; Oxford Knee Score, NA; Not Applicable. NRS pain; Numerical Rating Scale for pain during activity, 30s CST; 30-second Chair Stand Test, 40m FPWT; 40-meter Fast-Paced Walk Test, 10-step SCT; 10-step Stair Climb Test, * The 10-step SCT is scored in the opposite direction of the 30s CST and 40m FPWT (better performance is a lower score) therefore the hypothesised correlations is in the opposite direction.*

Table 5. Responsiveness.

	30s CST (change score)		40m FPWT (change score)		10-step SCT (change score)	
Predefined hypotheses	**Correlation Coefficient**	**Hypothesis Confirmed**	**Correlation Coefficient**	**Hypothesis Confirmed**	**Correlation Coefficient**	**Hypothesis Confirmed**
1. Moderate correlation with anchor question (≥ 0.4)	0.22	No	0.40	Yes	-0.25	No
2. Moderate correlation with change score NRS pain during activity (≤-0.4)	-0.20	No	-0.36	No	0.08	No
3. Moderate correlation with change score KOOS-PS (≤ -0.4)	-0.26	No	-0.28	No	0.27	No
4. Moderate correlation with change OKS (≥ 0.4)	0.22	No	0.43	Yes	-0.36	No
5. Correlation between change scores NRS pain and KOOS-PS is minimal 0.1 stronger than between NRS pain and performance-based test	0.56 / -0.20	Yes	0.56 / -0.36	Yes	-0.56 / 0.08	Yes
6. Correlation between change scores NRS pain and KOOS-PS is minimal 0.1 stronger than between KOOS-PS and performance-based test	-0.56 / -0.26	Yes	0.56 / -0.28	Yes	-0.56 / 0.27	Yes
7. Correlation between changes scores NRS pain and OKS minimal 0.1 stronger than between NRS pain and performance-based test	-0.70 / -0.20	Yes	-0.70 / -0.36	Yes	-0.70 /-0.08	Yes
8. Correlation between change scores NRS pain and OKS is minimal 0.1 stronger than between OKS and performance-based test	-0.70 / -0.22	Yes	-0.70 / 0.40	Yes	-0.70 / -0.36	Yes
Hypothesis confirmed	**4 / 8**	**50%**	**6 / 8**	**75%**	**4 / 8**	**50%**

KOOS-PS; Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form, OKS; Oxford Knee Score, NA; Not Applicable. NRS pain; Numerical Rating Scale for pain during activity, 30s CST; 30-second Chair Stand Test, 40m FPWT; 40-meter Fast-Paced Walk Test, 10-step SCT; 10-step Stair Climb Test.

References

1. Litwic A, Edwards MH, Dennison EM, Cooper C. Epidemiology and burden of osteoarthritis. *Br Med Bull*. 2013;105(1):185-199. doi:10.1093/bmb/lds038
2. Hoy D, Cross M, Smith E, et al. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis*. 2014;73(7):1323-1330. doi:10.1136/annrheumdis-2013-204763
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. *Osteoarthr Cartil*. 2008;16(2):137-162. doi:10.1016/j.joca.2007.12.013
4. 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):389-399. doi:10.1016/j.joca.2004.02.001
5. Reiman MP, Manske RC. The assessment of function: How is it measured? A clinical perspective. *J Man Manip Ther*. 2011;19(2):91-99. doi:10.1179/10669811 1X12973307659546
6. Terwee CB, Mokkink LB, Steultjens MPM, Dekker J. Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Rheumatology (Oxford)*. 2006;45(7):890-902. doi:10.1093/rheumatology/kei267
7. Mizner RL, Petterson SC, Clements KE, Zeni J a, Irrgang JJ, Snyder-Mackler L. Measuring functional improvement after total knee arthroplasty requires both performance-based and patient-report assessments: a longitudinal analysis of outcomes. *J Arthroplasty*. 2011;26(5):728-737. doi:10.1016/j.arth.2010.06.004
8. Stevens-Lapsley JE, Schenkman ML, Dayton MR. Comparison of self-reported knee injury and osteoarthritis outcome score to performance measures in patients after total knee arthroplasty. *PM R*. 2011;3(6):541-549; quiz 549.doi:10.1016/j.pmrj.2011.03.002
9. Stratford PW, Kennedy DM, Woodhouse LJ. Performance measures provide assessments of pain and function in people with advanced osteoarthritis of the hip or knee. *Phys Ther*. 2006;86(11):1489-1496. doi:10.2522/ptj.20060002
10. Dobson F, Hinman RS, Hall M, Terwee CB, Roos EM, Bennell KL. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis : a systematic review. *Osteoarthr Cartil*. 2012;20(12):1548-1562. doi:10.1016/j.joca.2012.08.015
11. Stratford PW, Kennedy DM. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. *J Clin Epidemiol*. 2006;59(2):160-167. doi:10.1016/j.jclinepi.2005.07.012
12. Hegedus EJ, Vidt ME, Tarara DT. The best combination of physical performance and self-report measures to capture function in three patient groups. *Phys Ther Rev*. 2014;19(3):196-203. doi:10.1179/1743288X13Y.0000000121
13. Dobson F, Hinman RS, Roos EM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. *Osteoarthr Cartil*. 2013;21(8):1042-1052.doi:10.1016/j.joca.2013.05.002
14. World Health Organization. *International Classification of Functioning, Disability and Health*.; 2001.
15. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. Cambridge: Cambridge University Press; 2011. doi:10.1017/CBO9780511996214
16. 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):539-549. doi:10.1007/s11136-010-9606-8

17. Gill S, McBurney H. Reliability of performance-based measures in people awaiting joint replacement surgery of the hip or knee. *Physiother Res Int*. 2008;13(3):141-152. doi:10.1002/pri.411
18. Kennedy DM, Stratford PW, Wessel J, Gollish JD, Penney D. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. *BMC Musculoskelet Disord*. 2005;6:3. doi:10.1186/1471-2474-6-3
19. Gustavo J. Almeida, PT M, Carolyn A. Schroeder B, Alexandra B. Gil, PT M, Kelley G, Fitzgerald, PT P, Sara R. Piva, PT P. Inter-Rater Reliability and Validity of the Stair Ascend/Descend Test in Individuals With Total Knee Arthroplasty. *Arch Phys Med Rehabil*. 2010;91(9):932–938. doi:10.1124/dmd.107.016501.CYP3A4-Mediated
20. Perruccio A V, Lohmander LS, 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. *Osteoarthr Cart*. 2008;16:542-550. doi:10.1016/j.joca.2007.12.014
21. Ruyssen-Witrand A, Fernandez-Lopez CJ, Gossec L, Anract P, Courpied JP, Dougados M. Psychometric properties of the OARSI/OMERACT osteoarthritis pain and functional impairment scales: ICOAP, KOOS-PS and HOOS-PS. *Clin Exp Rheumatol*. 2011;29(2):231-237.
22. de Groot IB, Favejee MM, Reijman M, Verhaar J a N, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. *Health Qual Life Outcomes*. 2008;6:16. doi:10.1186/1477-7525-6-16
23. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Outcome. *Arthritis Care Res (Hoboken)*. 2011;63 Suppl 1(November):S208-28. doi:10.1002/acr.20632
24. Haverkamp D, Breugem SJM, 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 Orthop*. 2005;76(3):347-352. doi:10.1080/00016470510030814
25. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343.
26. Conner-Spady BL, Marshall DA, Bohm E, et al. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. *Qual Life Res*. 2015;24(7):1775-1784. doi:10.1007/s11136-014-0910-6
27. Cibere J, Thorne A, Bellamy N, et al. Reliability of the Knee Examination in Osteoarthritis. *Arthritis rheum*. 2004;50(2):458-468. doi:10.1002/art.23310
28. Holstege MS, Lindeboom R, Lucas C. Preoperative quadriceps strength as a predictor for short-term functional outcome after total hip replacement. *Arch Phys Med Rehabil*. 2011;92(2):236-241. doi:10.1016/j.apmr.2010.10.015
29. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34-42. doi:10.1016/j.jclinepi.2006.03.012
30. Prinsen CAC, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" – a practical guideline. *Trials*. 2016;17(1):449. doi:10.1186/s13063-016-1555-2
31. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol*. 2010;10:22. doi:10.1186/1471-2288-10-22

32. Atkinson G NA. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sport Med*. 1998;26(4):217-238. doi:10.1016/j.patrec.2005.10.010
33. Nunally, Bernstein. *Psychometric Theory. 3rd Edition.*; 1994.
34. Terwee CB, van der Slikke RMA, van Lummel RC, Benink RJ, Meijers WGH, de Vet HCW. Self-reported physical functioning was more influenced by pain than performance-based physical functioning in knee-osteoarthritis patients. *J Clin Epidemiol*. 2006;59(7):724-731. doi:10.1016/j.jclinepi.2005.11.019
35. Bolink SAAN, Grimm B, Heyligers IC. Patient-reported outcome measures versus inertial performance-based outcome measures: A prospective study in patients undergoing primary total knee arthroplasty. *Knee*. 2015;22(6):618-623. doi:10.1016/j.knee.2015.04.002
36. Steultjens MPM, Dekker J, van Baar ME, Oostendorp R a. B, Bijlsma JWJ. Internal consistency and validity of an observational method for assessing disability in mobility in patients with osteoarthritis. *Arthritis Rheum*. 1999;12(1):19-25. doi:10.1002/1529-0131(199902)12:1<19::AID-ART4>3.0.CO;2-V
37. Piva SR, Fitzgerald GK, Irrgang JJ, Bouzubar F, Starz TW. Get Up and Go Test in Patients with Knee Osteoarthritis. *Arch Phys Med Rehabil*. 2004;85(2):284-289. doi:10.1016/j.apmr.2003.05.001
38. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes*. 2003;1(1):17. doi:10.1186/1477-7525-1-17
39. Conaghan PG, Emerton M, Tennant A. Internal construct validity of the Oxford knee scale: Evidence from Rasch measurement. *Arthritis Care Res*. 2007;57(8):1363-1367. doi:10.1002/art.23091
40. Davis AM, Perruccio A V, Canizares M, et al. Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. *Osteoarthr Cartil*. 2009;17(7):843-847. doi:10.1016/j.joca.2009.01.005
41. Ramkumar PN, Harris JD, Noble PC. Patient-reported outcome measures after total knee arthroplasty: a systematic review. *Bone Jt Res*. 2015;4(7):120-127. doi:10.1302/2046-3758.47.2000380
42. Mills K, Hunt M a, Ferber R. Biomechanical deviations during level walking associated with knee osteoarthritis: a systematic review and meta-analysis. *Arthritis Care Res (Hoboken)*. 2013;65(10):1643-1665. doi:10.1002/acr.22015

Abstract

Background and purpose

Improvement of physical function is one of the main treatment goals in severe hip osteoarthritis (OA) patients. The Osteoarthritis Research Society International (OARSI) has identified a core set of performance-based tests to assess the construct physical function: 30-s chair-stand test (30s CST), 4x10-meter fast-paced-walk test (40m FPWT) and a stair-climb test. Despite this recommendation, available evidence on the measurement properties is limited. We evaluated the reliability, validity and responsiveness of these performance-based measures in patients with hip OA scheduled for total hip arthroplasty (THA).

Patients and methods

Baseline and 12-month follow-up measurements were prospectively obtained in 90 end-stage hip OA patients that underwent THA. As there is no gold standard for comparison, the hypothesis testing method was used for construct validity and responsiveness analysis. A test can be assumed valid if ≥75% of predefined hypotheses are confirmed. A subgroup (n=30) underwent test-retest measurements for reliability analysis. The Oxford Hip Score, Hip injury and Osteoarthritis Outcome Score - Physical Function Short Form, pain during activity score and muscle strength were used as comparator instruments.

Results

Test-retest reliability was appropriate; Intraclass Correlation Coefficient values exceeded 0.70 for all 3 tests. None of the performance-based measures reached 75% hypothesis confirmation for the construct validity or responsiveness analysis.

Interpretation

The performance-based tests have good reliability in the assessment of physical function. Construct validity and responsiveness, using patient-reported measures and muscle strength as comparator instruments, could not be confirmed. Therefore, our findings do not justify their use for clinical practice.

chapter 3

Measurement properties of the OARSI core set of performance-based measures for hip osteoarthritis.

A prospective cohort study on reliability, construct validity and responsiveness in 90 hip osteoarthritis patients.

J.J. Tolk, R.P.A. Janssen, C.A.C. Prinsen, D.A.J.M. Latijnhouwers, M.C. van der Steen, S.M.A. Bierma-Zeinstra, M. Reijman

Acta Orthopaedica. 2019;90(1)15-20.

Introduction

Improvement of physical function is one of the main treatment goals of total hip arthroplasty (THA). Physical function can be assessed using patient-reported and performance-based outcome measurement instruments.[1] Because different domains of the construct physical function are measured, the methods are considered complementary and not competing.[1–3]

3 activities have been identified as most relevant for patients with hip OA: sit-to-stand movement, level walking and stair negotiation.[3] Impairment on these domains is classified as 'activity limitations' on the World Health Organisation International Classification of Functioning, Disability and Health (ICF).[4] The Osteoarthritis Research Society International (OARSI) has identified a set of performance-based tests to assess the construct physical function.[3,5] The core set consists of the 30-s chair-stand test (30s CST) for assessment of sit-to-stand movement, 4x10 meter fast-paced walk test (40m FPWT) for assessment of level walking and a stair-climb test to asses stair-negotiation.[3]

The validity and responsiveness of the OARSI core set have been challenged in knee OA patients,[6] but available evidence on the measurement properties in patients with hip OA is insufficient.[3,5] Measurement properties of a test should be confirmed in the population it is to be used, but the recommendation to use the specific tests included in the OARSI core set is based on expert opinion.[3,5] Therefore, before further implementation of the OARSI core set for hip OA patients can be considered, additional evidence on the measurement properties of these performance measures is essential.[5,7] We evaluated the reliability, validity and responsiveness after THA of the OARSI recommended performance-based measures, for measurement of physical function in patients with severe hip OA.

Patients and methods

We performed a prospective cohort study of patients indicated for THA to evaluate the measurement properties of the 30s CST, 40m FPWT and 10-step stair climb test (10-step SCT). The study was conducted following the COSMIN (COnsensus based Standards for the selection of health status Measurement INstruments) checklist.[8] The COSMIN checklist contains design requirements and preferred statistical methods for studies on measurement properties of health status measurement instruments.

Patient population

Patients were eligible for inclusion if they had unilateral symptomatic hip OA and were scheduled for primary THA. Patients with comorbidity leading to inability to perform the performance-based measures, insufficient knowledge of the Dutch language and inability to visit follow-up appointments were excluded. All patients in the *Máxima Medical Centre* meeting these criteria, and willing to participate, signed an informed consent form. The number of patients needed for the analysis was guided by the COSMIN standards [8,9]. We aimed to include ≥50 patients for construct validity and responsiveness analyses, and 30 patients for reliability analyses.

Study procedures

Patient characteristics measured at baseline were: sex, age, and BMI.

The assessment of performance-based measures and comparator instruments described below was made at baseline before surgery, and 12-months after THA. The standardized testing procedures were performed by a research nurse strictly according to the manual provided by the OARSI, with a fixed order of tests.[3]

Performance-based measures

30-s CST

The 30-s CST aims to quantify a patients performance on the activity 'sit-to-stand movement'.[3] From a sitting position, the patient stands up until hips and knees are fully extended, then completely back down. This is repeated for 30 seconds and each full cycle is counted as 1 chair stand.[3] A 43-cm high, straight back chair without armrests was used. For patients with hip OA, good reliability is reported with an Intraclass Correlation Coefficient (ICC) of 0.81 (0.63 - 0.91) and Standard Error of Measurement (SEM) of 1.27.[10] No reports on construct validity are available.

40m FPWT

The 40m FPWT is a test for performance on the activity short distance walking.[3] Participants are asked to walk as quickly but as safely as possible, without running, along a 10-meter walkway for a total distance of 40 meters. Walking speed is measured in meters / second (m/s). Use of a walking aid is allowed and recorded. Inter-rater reliability is reported to be good in patients with hip OA, with ICC of 0.95 (0.90 - 0.98) and SEM of 1.0 m/s.[10] There are no reports available on the construct validity.

Stair Climb Test

The OARSI included a stair climb test in the core set, but no specific measure is recommended.[3] We selected the 10-step stair climb test (10-step SCT), as the stair in the testing area had 10 steps with a step height of 19 cm. Patients were instructed to ascend and descend the flight of stairs as quickly as possible but in a safe manner. The time needed is recorded in seconds.[3] To our knowledge, there is no evidence available on measurement properties of the 10-step stair climb test or comparable stair climb tests in patients with hip OA.

Comparator instruments

We used a combination of comparator instruments, a specification of these instruments and their measurement properties can be found in a supplementary file. For measurement of physical function 2 joint-specific PROMs were used: the Hip injury and Osteoarthritis Outcome Score - Physical Function Short Form *(HOOS-PS)* [11], and the Oxford Hip Score (OHS).[12] The EuroQol 5D-3L (EQ-5D) was used as a measure of health-related quality of life.[13] Pain during activity was scored from 0 to 10 using a numerical rating scale (NRS pain).[14] At 12-months follow-up a 7-point Likert scale anchor question was scored for change in activities of daily living. Pre-operatively knee extensor and hip abductor strength of the affected leg was measured using a handheld dynamometer.[15,16]

Evaluation of the measurement properties and statistics

Reliability

Test-retest reliability refers to the extent to which scores for patients who have not changed are the same for repeated measurement over time. For this analysis, test-retest measurements of the 3 performance-based measures were obtained in a subset of the study population. 30 minutes of rest were allowed in between, to allow for full recovery during the resting interval. Performance on the activity under study can assumed to be stable over this testing period. ICC values for absolute agreement with corresponding 95% Confidence Intervals (CI) were calculated using a 2-way random model with absolute agreement. The threshold for an appropriate ICC is 0.70.[9,17] SEM and SDC were calculated as described by Atkinson.[18]

Construct validity

Construct validity refers to the degree in which the instruments under study measure the construct they aim to measure. This is the recommended method to assess validity when there is no 'Gold Standard' available, as is the case for the functional domains level walking, stair negotiation and sit-to-stand movement in hip OA. Before the start of the study, an expert panel formulated hypotheses on the expected relationships of performance-based measure scores with scores on the comparative instruments (Table 3).[19,20] Direction and magnitude of the expected results were stated. The expert panel consisted of an orthopaedic surgeon (RJ), orthopaedic resident and PhD candidate (JT), specialist in measurement property analysis (CP) and a methodologist (MR).
The hypotheses were based on the following predictions, we expected: a moderate correlation of the performance-based measures with PROMs and quadriceps strength, a stronger correlation of PROMs with pain scores than with the performance-based measures, a stronger correlation of the Performance-based measures with PROMs measuring functional outcome than with a PROM measuring general health, a stronger correlation of specific questions of the PROMs regarding walking, stair negotiation and sit-to-stand movement to their respective performance-based measure than to the total PROM score. Correlations on convergent hypothesis were expected to be at least moderate; ≥ 0.4 or ≤ -0.4. Divergent hypotheses were expected to have a poor correlation [≥ -0.39; ≤ 0.39]. Pearson or Spearman correlation coefficients were calculated, depending on normality of data distribution. Construct validity can be assumed adequate if at least 75% of the predefined hypotheses are confirmed.[9]

Responsiveness

Responsiveness refers to the ability of the instruments to detect change over time in the construct measured. In the absence of a gold standard, a construct approach is to be used. Hypotheses were formulated a priori by the expert panel, in a similar manner as for the construct validity analysis (Table 5).[9,19,20]

The hypotheses were formulated according to the following criteria: the anchor question would be moderately correlated to change in the performance-based measures scores (≥ 0.4 or ≤ -0.4) and the change in PROMs would be more correlated to pain, than to change in the performance-based measure scores. Pearson or Spearman correlation coefficients were calculated, depending on normality of data distribution. Adequate responsiveness can be assumed if minimally 75% of the predefined hypotheses are confirmed.[9]
SPSS statistics version 24.0 was used for the analyses (IBM Corporation).

Ethics, funding and potential conflicts of interest

The *Máxima Medical Centre* Medical Ethics Committee approved the study (registration code 2014-73). The authors declare that there are no conflicts of interest related to this article.

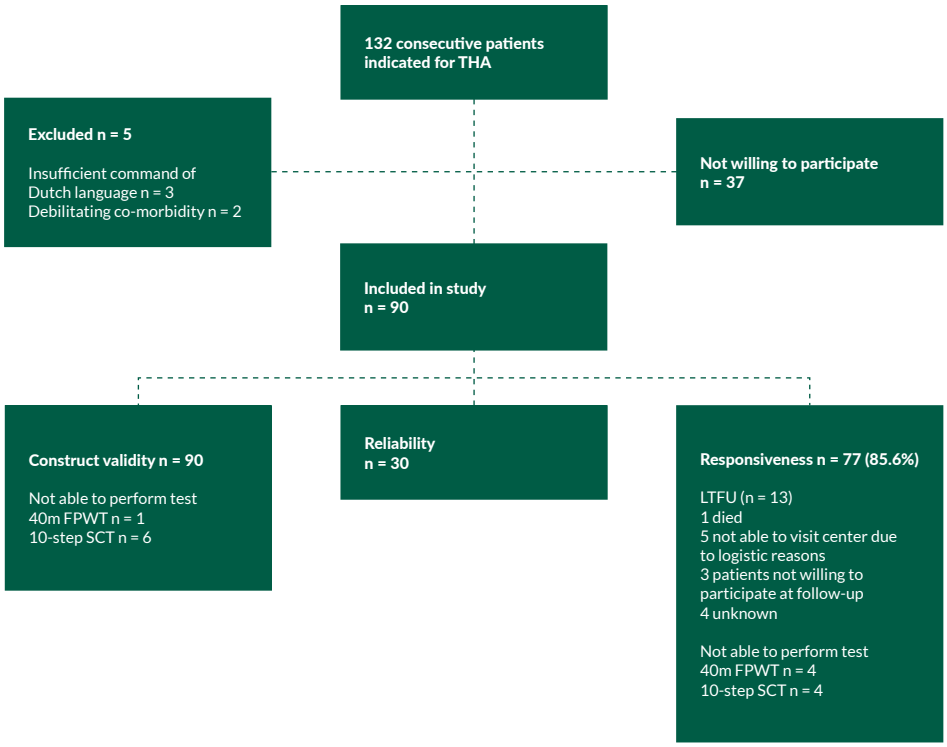


Figure 1. *Patients included in the analyses and lost to follow up. LTFU, lost to follow up.*

Results

Patient characteristics

In the period April to October 2015, 90 consecutive patients scheduled for arthroplasty because of hip OA were recruited (Table 1, Figure 1).

Measurement properties

a) Reliability analysis

30 randomly selected patients were enrolled in the test-retest study. Test-retest reliability was appropriate; Intraclass Correlation Coefficient values exceeded 0.70 for all 3 tests. (Table 2).

Table 1. Patient characteristics				
	Total cohort (n=90)		Reliability analysis cohort (n=30)	
Age, years	69.4	(9.5)	66.1	(9.4)
Gender female, n	61		22	
Side affected, right n	40		16	
BMI, kg/m2	27.0	(3.9)	26.0	(2.7)
Hip abductor strength, N	195.8	(7.8)	219.0	(7.9)
Knee extensor strength, N	133.5	(5.7)	136.9	(4.3)
Data are presented as mean and standard deviation between parentheses, or reported otherwise as mentioned.				

Table 2. Reliability analysis (n = 30)						
	Mean score baseline	**Mean retest score**	**Mean of Difference (baseline – retest score)**	**ICC (95% CI)**	**SEM**	**SDC**
30s CST (stands)	10.w1 (9.0-11.2)	10.9 (9.7-12.1)	-0.8 (-0.3 – -1.4)	0.86 (0.66 - 0.94)	0.99	2.7
40m FPWT (m/s)	1.32 (1.22-1.43)	1.33 (1.20-1.46)	-0.01 (-0.05 – 0.04)	0.94 (0.88 - 0.97)	0.08	0.22
10-step SCT (s)	14.2 (12.3-16.0)	14.1 (12.3-15.9)	-0.1 (-0.5 – 0.6)	0.96 (0.91 - 0.98)	1.06	2.9
ICC, Intraclass Correlation Coefficient; SEM, Standard Error of Measurement; SDC, Smallest Detectable Change. Data are presented as mean and 95% confidence interval between parentheses.						

b) Construct validity (hypothesis testing)

None of the 3 performance-based measures reached confirmation of 75% or more of the predefined hypotheses. 4/9 were confirmed for the 30s CST, 6/17 for the 40m FPWT and 6/17 for the 10-step SCT (Table 3).

c) Responsiveness

The mean score on the anchor question for change in activities of daily living (7-point Likert scale) at 12-month follow up was 6.2 (5.9-6.4), which represents 'much improvement' (Table 4) Results of the responsiveness analysis are presented in Table 5. For the 30s CST, 4/8 of the hypothesis were confirmed, for the 40m FPWT 5/8 and for the 10-step SCT 5/8.

Discussion

To our knowledge, this is the first thorough assessment of the measurement properties of the OARSI recommended core set of performance-based measures in patients with severe hip OA. The reliability analysis showed excellent test-retest reliability, which is in line

with previous reports.[10,21] Construct validity and responsiveness could not be confirmed. These findings are in accordance with recently published work on the OARSI core set of performance-based measures in knee OA patients.[6]

All 3 performance-based measures scored poorly on the construct validity and responsiveness analysis. One of the reasons is that almost all convergent hypotheses with PROMs measuring physical function were rejected. Although both methods aim to quantify related constructs, previous research has shown that PROMs assessing physical function do not measure the exact same domain as performance-based measures.[1–3] This potentially limits the strength of the conclusions that can be drawn from the present study. For example, PROMs are known to have a higher dependency on pain scores than performance based-measures.[2] When in the absence of a gold standard the construct approach is to be used, it is inherently so that there is a discrepancy between the test under study and the comparator instruments.[19] Furthermore, PROMs were not the only comparative instruments used, and hypotheses predicting a higher correlation of the performance-based measure scores with related construct compared to less related constructs were largely rejected as well. Therefore, in our opinion the conclusion on the construct validity and responsiveness should be interpreted broader than only showing the known discrepancy between PROMs and these measures.

As an alternative to the comparator instruments used for construct validity and responsiveness in the present study, 3-D motion analysis or inertia-based motion analysis could be used. These methods allow for a kinematic analysis in patients with hip OA, but their clinical relevance has not been defined.[22,23] Therefore, we believe these alternative methods are not suitable for comparison purposes in a clinical perspective. The comparative instruments used in the present study were considered the most suitable instruments available.

The findings on construct validity of the performance-based measures might be affected because impairment on the tested activities in daily living is not fully appreciated by merely timing the performance.[2,24] Although others claim good face validity for the core set of performance-based measures,[3,21] in our view this is not straightforward. For example, standing up and sitting down in rapid sequence, as measured by the 30s CST, is not really exemplary for stand-to-sit movement in daily life. Fewer repetitions on the test does not necessarily mean the quality of a sit-to-stand movement in daily living is more or less impaired. The same goes for walking speed and stair ascent, which does not directly represent more or less impairment. Merely timing the activity or counting repetitions cannot capture impairment caused by limping or joint instability nor avoidance of an activity in daily living.[24,25] This is a possible explanation why the construct validity could not be confirmed.

The responsiveness analysis showed that change in pain scores was strongly correlated to change in PROM scores, but not related to performance-based measure scores. Others have presented this low correlation with pain scores as a strength of performance-based measures, claiming this makes them more 'objective'.[3,5] In our opinion, it seems unlikely that the degree of pain during an activity would not influence performance in daily living.[25] Furthermore it has been shown that pain during activity does affect the quality of

movement, and impaired quality of movement is associated with lower perceived physical function.[24,26]Although pain reduction is not related to an increase in speed on the tested activities, the quality and manner of performance might improve [24,27], and patients might no longer avoid the activities [25]. These factors of physical performance are not grasped by the performance-based measures under study. The number of repetitions or speed scored on the performance-based measures might be of interest for research purposes, but to the authors' opinion actual change and perceived change need to be related to some degree for a test to be clinically relevant. Hypotheses in this regard were all rejected, contributing to the negative conclusion on the responsiveness of the OARSI core set of performance-based measures.

The strict adherence to the methodological criteria provided by COSMIN, are a strength of the present study.[8] Most previous reports on the measurement properties of the performance-based measures under study reported combined groups of hip and knee OA patients, resulting in heterogeneous populations.[21,28,29] The present study reports on an unselected, consecutive group of only end-stage hip OA patients. The results can therefore be considered more accurate and representative for this population.

The groups size for test-retest measurements was kept relatively small, to reduce the burden of repeated measurements for patients. As there is evidence from other studies showing similar results on reliability,[10,21,29] in our view it can be concluded that the performance-based measures under study have adequate test-retest reliability. The percentage of patients lost to follow-up for the responsiveness analysis was 14%. In our opinion, this can be considered acceptable, especially as the group of patients with incomplete data did not show systematic difference in baseline characteristics (Table 1).

In summary, the 30s CST, 40m FPWT and 10-step SCT have good reliability in the assessment of the domains sit-to-stand movement, walking short distances and stair negotiation in the construct physical function. Construct validity and responsiveness, using patient-reported measures and muscle strength as comparator instruments, could not be confirmed. Therefore, the present study does not justify their use for clinical practice in patients with severe hip OA.

Acknowledgements

We would like to sincerely thank C. van Doesburg, H. Kox, D. Latijnhouwers and M. Mariam for their work in administrative and testing procedures.

Supplementary file – specification of comparator instruments used.

Comparator instruments

HOOS-PS

The Hip injury and Osteoarthritis Outcome Score - Physical Function Short Form (HOOS-PS) is a 5-item PROM for measurement of the construct physical function. The HOOS-PS is scored on a 0 to 100 scale, 0 indicating no symptoms and 100 indicating extreme symptoms.11 The HOOS-PS has good construct validity and responsiveness in hip OA patients.11

OHS

The Oxford Hip Score (OHS) is a 12-item disease specific PROM for measurement of pain and function of the hip in relation to different activities of daily life. The total score ranges from 12 indicating no difficulties symptoms to 60 indicating most difficulties.12 The OHS has shown to be consistent, reliable, valid and sensitive to clinical change.12,30

EQ-5D

EuroQol 5D-3L (EQ-5D) is a standardized instrument developed as a measure of health-related quality of life.13 This PROM consists of a 5-question descriptive part and a visual analogue scale score (EQ-VAS) ranging from 0 to 100.13 From the 5-question part a sum score can be calculated, where 1 represents the best possible health state and lower scores represent worse health state.13 The EQ-5D has shown to be valid and reliable in hip OA patients.31

NRS pain

Pain during activity was scores using a numerical rating scale (NRS pain). Patients were asked to score pain during activity in the past week on an 11-point scale, the patients rate their pain during activity from 0 to 10. A score of 0 represented 'no pain' and a score of 10 represented 'worst imaginable pain'. Good reliability and responsiveness are reported for this NRS pain scale.14

Anchor question

At 12-months follow-up a 7-point Likert scale anchor question was scored for change in activities of daily living. The question 'how has your general daily functioning changed since the operation on your knee?' was scored from 1 (a lot worse) to 7 (very much improved).

Muscle strength

Strength of the knee extensors and hip abductors of the affected leg were tested for all subjects in the study. Maximal isometric knee extensor strength was measured in Newton (N) using a handheld dynamometer (HHD). In an upright sitting position, the HHD was positioned on the anterior aspect of the tibia, five cm proximal to the medial malleolus. A protective shin guard was used for patient comfort as well as standardization of HHD placement. Hip abductor strength was measured with subjects in supine position and with 5° of hip abduction. The HHD was positioned on the lateral femoral condyle and its position was held constant between trials to avoid changes in the resistance moment arm. For both muscle groups three consecutive measurements were obtained, the highest value was used for analysis. The HHD is a widely used instrument to measure knee extensor and hip abductor strength, with good reliability in OA patients. An ICC of 0.94 – 0.97 is reported.15,16

Table 3. Construct validity.						
	30s Chair Stand Test		40m Fast-Paced Walk Test		10-step Stair Climb Test	
Predefined hypotheses	**Spearman Correlation Coefficient**	**Hypothesis Confirmed**	**Spearman Correlation Coefficient**	**Hypothesis Confirmed**	**Spearman Correlation Coefficient** (* scored in opposite direction)	**Hypothesis Confirmed**
1. Moderate correlation with HOOS-PS (≤ -0.4) *	0.21	No	0.21	No	-0.24	No
2. Moderate correlation with OHS (≥ 0.4) *	0.45	Yes	0.34	No	-0.27	No
3. Moderate correlation with hip abductor strength (≥ 0.4) *	0.21	No	0.48	Yes	-0.44	Yes
4. Moderate correlation with Quadriceps strength (≥ 0.4) *	0.35	No	0.46	Yes	-0.53	Yes
5. Unrelated with EQ-5D [-0.39; 0.39]	0.38	Yes	0.31	Yes	0.34	Yes
6. Correlation with HOOS-PS is minimal 0.1 stronger than with EQ-5D	0.21 / 0.38	No	0.21 / 0.31	No	-0.24 / 0.34	No
7. Correlation with OHS is minimal 0.1 stronger than with EQ-5D	0.45 / 0.38	No	0.34 / 0.21	Yes	-0.27 / 0.34	No
8. 'Absolute' correlation between NRS pain and HOOS-PS is minimal 0.1 higher than between performance-based measure and NRS pain	- 0.53 / -0.19	Yes	-0.53 / -0.12	Yes	- 0.53 / 0.02	Yes
9. 'Absolute' correlation between NRS pain and OHS is minimal 0.1 higher than performance-based measure and NRS pain	-0.63 / -0.19	Yes	-0.63 / -0.12	Yes	-0.63 / 0.02	Yes
10. 'Absolute' correlation **40m FPWT** with HOOS-PS Question 4 is minimal 0.1 stronger than with HOOS-PS	NA		-0.12 / 0.21	No	NA	
11. 'Absolute' **40m FPWT** with HOOS-PS Question 4 is minimal 0.1 stronger than with OHS	NA		-0.12 / 0.34	No	NA	
12. 'Absolute' correlation **40m FPWT** with HOOS-PS Question 4 is minimal 0.1 higher than with EQ-5D Score	NA		-0.12 / 0.31	No	NA	
13. 'Absolute' correlation **40m FPWT** with OHS Question 4 is minimal 0.1 stronger than with HOOS-PS	NA		-0.12 / 0.21	No	NA	

14. 'Absolute' correlation **40m FPWT** with OHS Question 4 is minimal 0.1 stronger than with OHS	NA		-0.12 / 0.34	No	NA	
15. 'Absolute' correlation **40m FPWT** with OHS Question 4 is minimal 0.1 stronger than with EQ-5D Score	NA		-0.12 / 0.31	No	NA	
16. Moderate correlation **40m FPWT** with EQ-5D Question 1 (≤ -0.4)	NA		-0.36	No	NA	
17. Moderate correlation **40m FPWT** with OHS Question 4 (≤ -0.4)	NA		-0.12	No	NA	
18. 'Absolute' correlation **10-step SCT** with OHS Question 6 is minimal 0.1 stronger than with HOOS-PS	NA		NA		0.31 / -0.24	No
19. 'Absolute' correlation **10-step SCT** with OHS Question 6 is minimal 0.1 stronger than with OHS	NA		NA		0.31 / -0.27	No
20. 'Absolute' correlation **10-step SCT** with OHS Question 6 is minimal 0.1 stronger than with EQ-5D	NA		NA		0.31 / -0.31	No
21. Moderate correlation **10-step SCT** with OHS Question 6 (≤-0.4)	NA		NA		0.31	No
22. 'Absolute' correlation **10-step SCT** with HOOS-PS Question 1 is minimal 0.1 stronger than with HOOS-PS	NA		NA		0.34 / -0.24	Yes
23. 'Absolute' correlation **10-step SCT** with HOOS-PS question 1 is minimal 0.1 stronger than with OHS	NA		NA		0.34 / -0.27	No
24. 'Absolute' correlation **10-step SCT** with HOOS-PS question 1 is minimal 0.1 stronger than with EQ-5D	NA		NA		0.34 / -0.31	No
25. Moderate correlation **10-step SCT** with HOOS-PS question 1 (≤-0.4)	NA		NA		0.34	No
Hypothesis confirmed	**4/9**		**6/17**		**6/17**	

*NA; Not Applicable. * The 10 step SCT is scored in the opposite direction of the 30s CST and 40m FPWT (better performance is a lower score) therefore the hypothesised correlations are in the opposite direction.*

Table 4. Performance-based measures and PROM scores before and after THA			
	Baseline	**12-month follow-up**	**p-value**
30s CST (stands)	9.3 (8.5-10.2)	12.0 (11.2-12.9)	<0.001
40m FPWT (m/s)	1.26 (1.17-1.34)	1.34 (1.26-1.42)	<0.001
Use of assistive device during 40m FPWT (patients, n)	8	2	0.057
10-step SCT (seconds)	17.9 (15.3-20.4)	14.5 (12.9-16.2)	<0.001
Use of handrail 10-step SCT (patients, n)	41	28	0.047
HOOS-PS score	48.0 (44.3-51.9)	21.7 (19.8-26.2)	<0.001
OHS	23.6 (21.9-25.7)	41.8 (40.5-43.2)	<0.001
EQ5D	0.51 (0.43-0.57)	0.83 (0.79-0.86)	<0.001
EQ VAS	64.8 (59.6-70.0)	76.1 (71.5-80.7)	0.001
NRS pain	6.8 (6.5-7.3)	1.5 (1.7)	<0.001
Anchor question (patients, n)			
Very much improvement		34	
Much improvement		33	
A little improvement		5	
Unchanged		1	
A little worse		0	
Much worse		4	
Very much worse		0	
Data are presented as mean and 95% confidence interval between parentheses, or reported otherwise as mentioned.			

Table 5. Responsiveness.						
	30s Chair Stand Test (change score)		40m Fast-Paced Walk Test (change score)		10-step Stair Climb Test (change score)	
Predefined hypotheses	**Spearman Correlation Coefficient**	**Hypothesis Confirmed**	**Spearman Correlation Coefficient**	**Hypothesis Confirmed**	**Spearman Correlation Coefficient**	**Hypothesis Confirmed**
1. Moderate correlation with anchor question (≥ 0.4)	0.37	No	0.28	No	-0.18	No
2. Moderate correlation with change score NRS pain during activity (≤-0.4)	-0.04	No	-0.13	No	0.14	No
3. Moderate correlation with change score HOOS-PS (≤ -0.4)	0.30	No	0.21	No	-0.35	No
4. Moderate correlation with change OHS (≥ 0.4)	0.23	No	0.27	No	-0.26	No
5. Correlation between change scores NRS pain and HOOS-PS is minimal 0.1 stronger than between NRS pain and performance-based test	-0.45 / -0.04	Yes	-0.45 / -0.13	Yes	-0.45 / -0.18	Yes
6. Correlation between change scores NRS pain and HOOS-PS is minimal 0.1 stronger than between HOOS-PS and performance-based test	-0.45 / 0.30	Yes	-0.45 / 0.21	Yes	-0.45 / -0.35	Yes
7. Correlation between changes scores NRS pain and OHS minimal 0.1 stronger than between NRS pain and performance-based test	-0.66 / -0.04	Yes	-0.66 / -0.13	Yes	-0.66 / -0.18	Yes
8. Correlation between change scores NRS pain and OHS is minimal 0.1 stronger than between OHS and performance-based test	-0.66 / 0.23	Yes	-0.66 / 0.27	Yes	-0.66 / -0.26	Yes
Hypothesis confirmed	**4/8**		**4/8**		**5/8**	

References

1. Reiman MP, Manske RC. The assessment of function: How is it measured? A clinical perspective. *J Man Manip Ther*. 2011;19(2):91-99. doi:10.1179/10669811 1X12973307659546
2. Stratford PW, Kennedy DM. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. *J Clin Epidemiol*. 2006;59(2):160-167. doi:10.1016/j. jclinepi.2005.07.012
3. Dobson F, Hinman RS, Roos EM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. *Osteoarthr Cartil*. 2013;21(8):1042-1052. doi:10.1016/j.joca.2013.05.002
4. World Health Organization. *International Classification of Functioning, Disability and Health*.; 2001.
5. Dobson F, Hinman RS, Hall M, Terwee CB, Roos EM, Bennell KL. Measurement properties of performance-based measures to assess physical function in hip and knee osteoarthritis : a systematic review. *Osteoarthr Cartil*. 2012;20(12):1548-1562. doi:10.1016/j. joca.2012.08.015
6. Tolk JJ, Janssen RPA, Prinsen CAC, et al. The OARSI core set of performance-based measures for knee osteoarthritis is reliable but not valid and responsive. *Knee Surgery, Sport Traumatol Arthrosc*. 2019;27(9):2898-2909. doi:10.1007/s00167-017-4789-y
7. Terwee CB, Mokkink LB, Steultjens MPM, Dekker J. Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Rheumatology (Oxford)*. 2006;45(7):890-902. doi:10.1093/rheumatology/kei267
8. 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):539-549. doi:10.1007/ s11136-010-9606-8
9. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34-42. doi:10.1016/j. jclinepi.2006.03.012
10. Wright A a, Cook CE, Baxter GD, Dockerty JD, Abbott JH. A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. *J Orthop Sports Phys Ther*. 2011;41(5):319-327. doi:10.2519/jospt.2011.3515
11. Davis AM, Perruccio A V, Canizares M, et al. Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. *Osteoarthr Cartil*. 2009;17(7):843-847. doi:10.1016/j.joca.2009.01.005
12. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *Bone Joint J*. 1996;78(2):185-190. doi:10.1302/0301-620X.80B1.7859
13. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343.
14. Ruyssen-Witrand A, Fernandez-Lopez CJ, Gossec L, Anract P, Courpied JP, Dougados M. Psychometric properties of the OARSI/OMERACT osteoarthritis pain and functional impairment scales: ICOAP, KOOS-PS and HOOS-PS. *Clin Exp Rheumatol*. 2011;29(2):231-237.
15. Holstege MS, Lindeboom R, Lucas C. Preoperative quadriceps strength as a predictor for short-term functional outcome after total hip replacement. *Arch Phys Med Rehabil*.

2011;92(2):236-241. doi:10.1016/j.apmr.2010.10.015
16. Zeni J, Abujaber S, Pozzi F, Raisis L. Relationship Between Strength, Pain, and Different Measures of Functional Ability in Patients With End-Stage Hip Osteoarthritis. *Arthritis Care Res (Hoboken)*. 2014;66(10):1506-1512. doi:10.1002/acr.22329
17. Prinsen CAC, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" – a practical guideline. *Trials*. 2016;17(1):449. doi:10.1186/s13063-016-1555-2
18. Atkinson G NA. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sport Med*. 1998;26(4):217-238. doi:10.1016/j.patrec.2005.10.010
19. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. Cambridge: Cambridge University Press; 2011. doi:10.1017/CBO9780511996214
20. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol*. 2010;10:22. doi:10.1186/1471-2288-10-22
21. Dobson F, Hinman RS, Hall M, et al. Reliability and measurement error of the Osteoarthritis Research Society International (OARSI) recommended performance-based tests of physical function in people with hip and knee osteoarthritis. *Osteoarthr Cartil*. 2017:6-10. doi:10.1016/j.joca.2017.06.006
22. Bolink SAAN, Lenguerrand E, Brunton LR, et al. Assessment of physical function following total hip arthroplasty: Inertial sensor based gait analysis is supplementary to patient-reported outcome measures. *Clin Biomech*. 2016;32:171-179. doi:10.1016/j.clinbiomech.2015.11.014
23. Kolk S, Minten MJM, Van Bon GEA, et al. Gait and gait-related activities of daily living after total hip arthroplasty: A systematic review. *Clin Biomech*. 2014;29(6):705-718. doi:10.1016/j.clinbiomech.2014.05.008
24. Steultjens MPM, Dekker J, van Baar ME, Oostendorp R a. B, Bijlsma JWJ. Internal consistency and validity of an observational method for assessing disability in mobility in patients with osteoarthritis. *Arthritis Rheum*. 1999;12(1):19-25. doi:10.1002/1529-0131(199902)12:1<19::AID-ART4>3.0.CO;2-V
25. Holla JFM, Sanchez-Ramirez DC, van der Leeden M, et al. The avoidance model in knee and hip osteoarthritis: a systematic review of the evidence. *J Behav Med*. 2014;37(6):1226-1241. doi:10.1007/s10865-014-9571-8
26. Rosenlund S, Holsgaard-Larsen A, Overgaard S, Jensen C. The Gait Deviation Index is associated with hip muscle strength and patient-reported outcome in patients with severe hip osteoarthritis - A cross-sectional study. *PLoS One*. 2016;11(4):1-13. doi:10.1371/journal.pone.0153177
27. Steultjens MP, Roorda LD, Dekker J, Bijlsma JW. Responsiveness of observational and self-report methods for assessing disability in mobility in patients with osteoarthritis. *Arthritis Rheum*. 2001;45(15):56-61. doi:10.1002/1529-0131(200102)45:1<56::AID-ANR84>3.0.CO;2-Y
28. Gill S, McBurney H. Reliability of performance-based measures in people awaiting joint replacement surgery of the hip or knee. *Physiother Res Int*. 2008;13(3):141-152. doi:10.1002/pri.411
29. Kennedy DM, Stratford PW, Wessel J, Gollish JD, Penney D. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. *BMC Musculoskelet Disord*. 2005;6:3. doi:10.1186/1471-2474-6-3
30. Gosens T, Hoefnagels N, de Vet R, et al. The "Oxford Heup Score": the translation and validation of a questionnaire into Dutch to evaluate the results of total hip arthroplasty. *Acta*

Orthop. 2005;76(2):204-211.

31. Conner-Spady BL, Marshall DA, Bohm E, et al. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. *Qual Life Res*. 2015;24(7):1775-1784. doi:10.1007/s11136-014-0910-6

Expectations of treatment result

part 02

Abstract

Introduction

The rate of satisfaction after total knee arthroplasty (TKA) is consistently reported around 80%, leaving 1 in 5 patients unsatisfied to some extent. Fulfilment of expectations is reported as the strongest predictor of treatment satisfaction. In this study we aimed to evaluate what Dutch orthopaedic surgeons assume are realistic expectations for recovery 1 year after TKA.

Methods

We invited the members of the Dutch Knee Society to fill out a web-based questionnaire. For expectation measurement the validated Dutch version of the Hospital for Special Surgery (HSS) knee replacement expectations survey was used.

Results

150 invitations were successfully sent. 84 orthopaedic surgeons responded (56%). The overall HSS knee replacement expectation score was 66.0 (SD 14.0) on a 0-100 scale. Most improvement was predicted for the items "pain relief" and "walking short distances". Expectations related to patients' ability to kneel or squat after TKA were scored poorly.

Conclusion

To the opinion of the members of the Dutch Knee Society, after TKA improvement can be expected in domains of pain, function, activities and psychological wellbeing. Return to normal is not likely to occur, especially in demanding physical activities.

chapter 4

Total knee arthroplasty, what to expect?

A survey of the members of the Dutch Knee Society on long-term recovery after total knee arthroplasty.

J.J. Tolk, M.C. van der Steen, R.P.A. Janssen, M. Reijman

Journal of Knee Surgery. 2017 Jul;30(6):612-616.

Introduction

Total knee arthroplasty (TKA) is a frequently performed procedure. The numbers of knee replacements in the Netherlands are increasing, with 26.754 performed in 2014 compared to 20.558 in 2010, according to the 2014 Annual Report of the Dutch Arthroplasty Register.[1] It is a relatively safe and cost-effective surgical intervention for patients with end-stage osteoarthritis and in general the results after TKA are good.[2] Treatment success can be measured by different methods. In addition to more traditionally used methods as prosthesis alignment, postoperative range of motion and complication percentage, nowadays patient reported outcome measures (PROMs), functional performance measures and patient satisfaction are increasingly considered as criteria of treatment success.[3–5] The rate of satisfaction is consistently reported around 80%, leaving 1 in 5 patients unsatisfied to some extent after their knee surgery.[3]

Multiple factors influence postoperative patient satisfaction after TKA. The fulfilment of preoperative expectations has been reported as one of the main determinants of postoperative satisfaction.[6–8] In a study by Hamilton et al. analysing factors affecting postoperative satisfaction after primary TKA in 2247 patients, the main predictor of satisfaction was 'meeting preoperative expectations'.[6] Bourne et al. found similar results in a study of 1703 primary TKA, where patients with expectations that were not met, were at 10.7x greater risk to be dissatisfied with treatment outcome.[9] Fulfilment of expectations is reported as the strongest predictor of treatment satisfaction, with more influence than pain relief, postoperative complications and pre- or postoperative physical status.[6,9] In this light expectation management in TKA patients, resulting in more realistic postoperative expectations, is thought advantageous to achieve optimal patient satisfaction.

Patients' recovery expectations are based on different sources, such as previous experiences, social network influences and health professional information transfer.[10,11] In the Netherlands, the orthopaedic surgeon and his/her team are obliged to include pre-operative education and expectation management as part of the informed consent process.[12] Previous research has shown that structured pre-operative education can reduce the rate of discordance between patients' and surgeons' expectations.[10,13] In this study we aimed to evaluate what Dutch orthopaedic surgeons assume are realistic expectations for long-term recovery after total knee arthroplasty.

Material and methods

To evaluate the opinion of Dutch orthopaedic surgeons, we invited all members of the Dutch Knee Society (DKS) to fill out a web-based questionnaire. The survey was conducted and reported according to the 'Best Practices for Survey Research Reports' recommendations.[14]

For expectation measurement the Dutch version of the Hospital for Special Surgery (HSS) knee replacement expectations survey was used.[15] This is a 19-item self-administered survey, measuring expectations in domains of pain, function, activities and psychological wellbeing.[16] The Dutch version of the HSS Knee Replacement Expectations Survey addresses 'probability based expectations' as defined by Kravitz [17], within the construct 'outcome

expectations' according to the preliminary framework by Crow et al [18]. The survey is reliable, validated and considered to be a high-quality expectation assessment instrument [16,19], containing the most important questions regarding long term recovery after TKA [16,20]. Response options are "complete improvement or back to normal" (4 points), "a lot of improvement" (3 points), "a moderate amount of improvement" (2 points), "a little improvement" (1 point), or "I do not have this expectation" (0 points).[10] Scores for each question can be presented as a mean score, and a total score for the survey can be calculated. The summed raw score ranges from 0 to 76 and the transformed score [= (raw score/76) · 100] ranges from 0 to 100. Higher scores indicate expecting more improvement for more items.[10,16]

All 152 members of the Dutch Knee Society (DKS) were invited to fill out this questionnaire according to what, on the basis of their expertise, an average patient with primary knee osteoarthritis can expect of his or her TKA. This time period was chosen because clinically important improvement occur at least until 12 month postoperatively.[21]

The invitation for participation was sent by e-mail at the end of November 2015, with an explication of the goal of the study and a link to the web-based survey. The survey was conducted using Research Manager (Cloud9 Software, version 5.7.1.0), which allows for automatized and anonymised administration of the questionnaires. Reminders were sent after 2 and 6 weeks. After 10 weeks the invitation expired and the survey could no longer be filled out.

Results

A total of 152 surveys were sent to all members of the DKS. Two DKS members could not be reached due to an invalid e-mail address. Of the 150 invitations successfully sent, we received 84 responses. Of these 84 questionnaires, 2 were only partially completed.

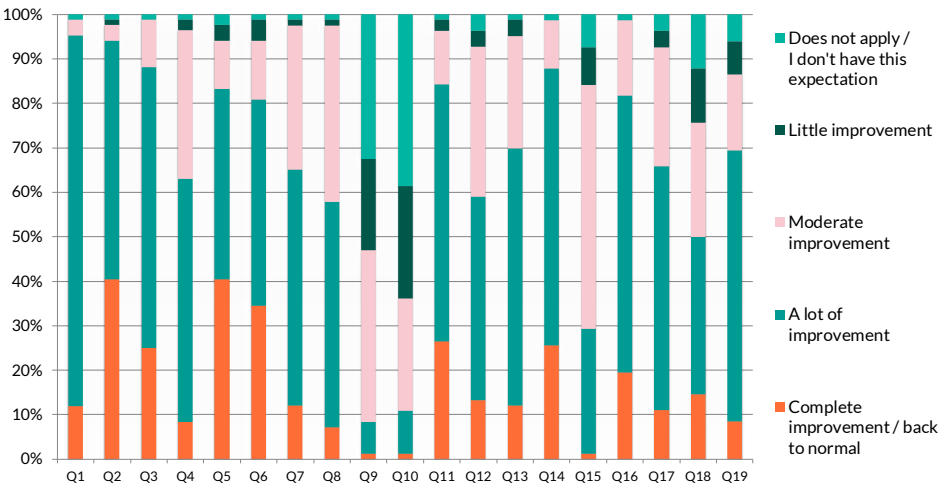


Figure 3. *Response to the Hospital for Special Surgery Knee Replacement Expectations Survey, questions numbered in the same order as described in Table 1.*

Table 1. Answers to the HSS Knee Replacement Expectations Survey.		
	Mean score (SD)	No. of respondents (%)
Q1. Relieve pain	3.1 (0.5)	84 (56)
Q2. Improve ability to walk short distance (indoors, 1 block)	3.3 (0.7)	84 (56)
Q3. Improve ability to walk medium distance (take a walk, less than 1 mile)	3.1 (0.7)	84 (56)
Q4. Improve ability to walk long distance (more than 1 mile)	2.7 (0.7)	84 (56)
Q5. Remove the need for a cane, crutch or walker	3.2 (0.9)	84 (56)
Q6. Make knee or leg straight	3.1 (0.9)	84 (56)
Q7. Improve ability to go up stairs	2.7 (0.7)	83 (55)
Q8. Improve ability to go down stairs	2.6 (0.7)	83 (55)
Q9. Improve ability to kneel	1.2 (1.0)	83 (55)
Q10. Improve ability to squat	1.1 (1.0)	83 (55)
Q11. Improve ability to use public transportation or drive	3.1 (0.8)	83 (55)
Q12. Be employed for monetary reimbursement	2.6 (0.9)	83 (55)
Q13. Improve ability to participate in recreational activities (for example dancing, pleasure travel)	2.8 (0.8)	83 (55)
Q14. Improve ability to perform daily activities (for example daily routines, household chores)	3.1 (0.7)	82 (55)
Q15. Improve ability to exercise or participate in sports	2.1 (0.8)	82 (55)
Q16. Improve ability to change position (for example go from sitting to standing or from standing to sitting)	3.0 (0.7)	82 (55)
Q17. Improve ability to interact with others (take care of someone, play with children)	2.7 (0.9)	82 (55)
Q18. Improve sexual activity	2.3 (1.2)	82 (55)
Q19. Improve psychological well-being	2.6 (0.9)	82 (55)

Overall a response rate of 56% was achieved. Of the incomplete questionnaires, only the questions that were filled out were included in the analysis.

Table 1 shows the response rate for the individual items and mean scores. The mean overall survey score was 66.0 (SD 14.0) on a 0-100 scale. Distribution of DKS members expectations for the different items of the survey are shown in Figure 1.

Discussion

To our knowledge this is the first report on what orthopaedic surgeons think to be realistic expectations for the average primary TKA patient at 1-year follow-up.

Overall the consulted orthopaedic surgeons had positive expectations of the outcome after TKA. For 8 out of 18 question, 80% or more of the surgeons responded that a lot or even total improvement can be expected. Most improvement was predicted for the items "pain relief" and "walking short distances", where more than 95% of respondents expected a large or total improvement. Pain relief is considered one of the main goals in knee arthroplasty by patients and surgeons [16,22], and significant pain relief is achieved in most patients after TKA [7,23]. In previous reports, much less or less pain is described in 93% of TKA patients at 12-month follow-up.[7] Nevertheless this pain relief process is gradual, variable and residual pain is common.[23,24] An unfavourable long-term pain outcome is reported in approximately 20% of patients [25], and only 43% of patients reported to be completely without pain at 1 year follow-up [24]. These reports are in line with the respondents' opinion in the present study that pain relief can be expected, but some residual pain is likely.

The orthopaedic surgeons rarely scored "complete improvement / back to normal" for the majority of the items. Only the items "ability to walk short distances" and "removal of a walking aid" scored reasonably for full recovery, although only 40% of surgeons respectively think full recovery is to be expected for these items. Previous literature has shown a considerable effect of TKA on short distance walking speed.[26,27] It has even been suggested that walking speed normalizes after knee arthroplasty when compared to healthy peers.[26] Walking distance has been reported to improve as well. In a prospective study in 102 TKA patients by Nilsdotter et al. 60% of TKA patients were able to walk more than 1 km without any complaints at 12 months follow up, whereas pre-operatively this was less than 5% of patients.[7] An unlimited walking distance was reported in approximately 40% of patients after TKA [7]. However, 20% of patients was reported to walk less than 1 km and 10% could only walk in their house.[7] A walking aid is seldom necessary after knee arthroplasty, less than 5% needed crutches at long term follow-up.[7] These findings in the literature reflect the surgeons' opinion that for "walking short distances" a large improvement can be expected. For longer distances limitations are likely to remain, although a walking aid is probably no longer necessary.

Another reflection of the generally positive view of outcome after TKA is the fact that the answer "this expectation does not apply to me / I do not have this expectation" is scored least overall. Only on 3 questions more than 10% of the orthopaedic surgeons chose this option. Especially the ability to kneel or squat after TKA is scored poorly, with respectively 33% and 39% of respondents expecting no improvement at all. In previous studies reporting data of kneeling after TKA, low recovery rates were reported with 72 - 82% of the patients being unable to kneel without some knee symptoms after TKA.[28,29] Furthermore 47 – 60 % reported moderate to severe difficulty on kneeling [28–30], and about 30% of patients was unable to kneel [29]. For squatting similar results are reported with 75 - 86% of patients unable to squat without some knee symptoms [28,29], 42 - 59 % had moderate to severe difficulty squatting [28,29], and about 25% of patients were unable to squat at all [28,29]. This is in agreement with the findings of the present study that little to no improvement in kneeling and squatting can be expected.

For most questions, the consulted surgeons' opinions did not show high variance, with standard deviation ≤1.0 point on a 5-point scale. An exception is question 18, "improve

sexual activity" for which the surgeons' responses are more variable. In literature resuming sexual activity after TKA is described in 85-94% of patients.[31,32] Not being able to engage in sexual activity due to the TKA is reported in only 1.6% of patients [32], but up to 14% reported their artificial knee to cause moderate to severe difficulty in sexual activities.[28] Frequency and ability to engage in sexual activity are reported inconsistently. Nunley et al. reported an increase in sexual frequency in up to 25% of patients due to less pain and greater mobility [32], while Klit et al. reported an increase in sexual frequency in only 1 out of 62 TKA patients, and a decrease in frequency in 32% [31]. Overall, sexual activity seems possible after TKA, but an actual improvement in ability or frequency should not be anticipated. A possible explanation for the variability in answers of the orthopaedic surgeons on this question might be that 'the average patient' in this regard is less well defined, and what is to be expected for this item maybe more dependent on other factors than the knee replacement. Another possibility is that expectation and outcome for this specific subject are not often discussed by orthopaedic surgeons [33], possibly due to embarrassment of the surgeon or the patient. This might result in less knowledge about what patients should expect, or outcomes that are commonly achieved. All of these points might have led to the variability in answers in the present study.

The overall survey score was 66. The previously reported HSS knee replacement expectation score in patients for probability-based expectations, are generally higher. Mancuso et al. reported an overall score of 85 in 143 patients awaiting TKA in a high volume orthopaedic centre in the United States.[10] Neuprez et al. reported the expectations in 61 patients scheduled for TKA in public and private teaching hospitals in Belgium, and found a mean score of 765.[34] As these reports are not from the same cohort as the consulted DKS members report their expectations on, direct comparison should be done with caution. The discordance is in line though with the report by Ghomrawi et al. that patient-surgeon disagreement on expectations is common, with patients generally having higher expectations.[13] Considering this discrepancy in the authors' opinion modification of patient expectations, resulting in more realistic pre-operative expectations, might be an effective an intervention to achieve higher patient satisfaction.

There are some limitations of the present study. There was a response rate of 56%. Although there is no general standard for a minimum acceptable response rate, general consensus is that at least half of the sample should have completed the survey instrument.[14] Because measurements were obtained anonymously, no comparison can be made between the DKS members that did complete the survey and those who did not. This may have led to a response bias, however the relatively high response rate lessens this risk of bias.

A factor that might have influenced the way the survey has been completed is the fact that 'the average patient' might not be the same for all respondents. Different orthopaedic surgeons might not have the same 'average' patients, as their patient population may differ e.g. university hospital, general hospital or private clinic. In contrast the inclusion of all types of orthopaedic practices makes the results of this survey more generalizable for what knee arthroplasty surgeons in the Netherlands overall think are realistic expectations for an average primary TKA patient. Because of the anonymity of the survey, stratification according to the professional setting of the respondent could not be made.

One of the strengths of the present study is the clear statement of the construct measured and theoretical framework it derives from. Use of high quality measurement instruments and interpretation according to the appropriate theoretical framework should be obvious, but unfortunately in the patient expectation field poor measurement methods are often used.[35,36] In the present study, the probability-based version of the HSS knee replacement expectations survey was used.[10] This surveys' clinimetric properties have been appropriately validated and this survey measures the desired construct; probability-based outcome expectations.[10,15,16]

Conclusion

Overall good effect for long-term recovery after TKA can be expected, according to the opinion of the members of the Dutch Knee Society. Improvement can be expected in domains of pain, function, activities and psychological wellbeing but return to normal is not likely to occur with limitations predominantly in demanding physical activities.

References

1. Hoy D, Cross M, Smith E, et al. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis*. 2014;73(7):1323-1330. doi:10.1136/annrheumdis-2013-204763
2. 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):148-152. doi:10.1302/0301-620x.95b11.32767
3. 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. doi:10.1016/j.arth.2017.07.021
4. Noble PC, Conditt M a, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452(452):35-43. doi:10.1097/01.blo.0000238825.63648.1e
5. Mancuso C, Graziano S, Briskie L, et al. Randomized trials to modify patients' preoperative expectations of hip and knee arthroplasties. *Clin Orthop Relat Res*. 2008;466(2):424-431. doi:10.1007/s11999-007-0052-z
6. Tolk JJ, van der Steen M, Janssen RPA, Reijman M. Total knee arthroplasty: what to expect ? A survey of the members of the Dutch Knee Society on long-term recovery after total knee arthroplasty. *J Knee Surg*. 2017;30(6):612-616. doi:10.1055/s-0036-1593868
7. Tolk JJ, Janssen RPA, Haanstra TM, Bierma-zeinstra SMA, Reijman M. The EKSPECT study : the influence of Expectation modification in Knee arthroplasty on Satisfaction of PatiEnts : study protocol for a randomized Controlled Trial. *Trials*. 2018;19(437).
8. Jain D, Nguyen L-CL, Bendich I, et al. Higher patient expectations predict higher patient-reported outcomes, but not satisfaction, in total knee arthroplasty patients: a prospective multicenter study. *J Arthroplasty*. 2017;32(9S):S166-S170. doi:10.1016/j.arth.2017.01.008
9. Hepinstall MS, Rutledge JR, Bornstein LJ, Mazumdar M, Westrich GH. Factors that impact expectations before total knee arthroplasty. *J Arthroplasty*. 2011;26(6):870-876. doi:10.1016/j.arth.2010.09.010
10. Smith EJ, Soon V-L, Boyd A, McAllister J, Deakin A, Sarungi M. What do Scottish Patients Expect of Their Total Knee Arthroplasty? *J Arthroplasty*. 2016;31(4):786-792. doi:10.1016/j.arth.2015.10.021
11. Bjorgul K, Novicoff WM, Saleh KJ. Evaluating comorbidities in total hip and knee arthroplasty: available instruments. *J Orthop Traumatol*. 2010;11:203-209. doi:10.1007/s10195-010-0115-x
12. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16(4):494-502. doi:10.1136/ard.16.4.494
13. van den Akker-Scheek I, van Raay JJ a M, Reininga IHF, Bulstra SK, Zijlstra W, Stevens M. Reliability and concurrent validity of the Dutch hip and knee replacement expectations surveys. *BMC Musculoskelet Disord*. 2010;11(1):242. doi:10.1186/1471-2474-11-242
14. Mancuso CA, Sculco TP, Wickiewicz TL, Jones EC, Robbins L, Warren RF. Patients' expectations of knee surgery. *J Bone Jt Surg Am*. 2001;83-A(7):1005-1012.
15. Zywiel MG, Mahomed A, Gandhi R, Perruccio A V., Mahomed NN. Measuring expectations in orthopaedic surgery: A systematic review. *Clin Orthop Relat Res*. 2013;471(11):3446-3456. doi:10.1007/s11999-013-3013-8
16. de Groot IB, Favejee MM, Reijman M, Verhaar J a N, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. *Health Qual Life Outcomes*. 2008;6:16. doi:10.1186/1477-7525-6-16
17. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function:

International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Outcome. *Arthritis Care Res (Hoboken)*. 2011;63 Suppl 1(November):S208-28. doi:10.1002/acr.20632

18. Haverkamp D, Breugem SJM, 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 Orthop*. 2005;76(3):347-352. doi:10.1080/00016470510030814
19. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343.
20. Conner-Spady BL, Marshall DA, Bohm E, et al. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. *Qual Life Res*. 2015;24(7):1775-1784. doi:10.1007/s11136-014-0910-6
21. Ruyssen-Witrand A, Fernandez-Lopez CJ, Gossec L, Anract P, Courpied JP, Dougados M. Psychometric properties of the OARSI/OMERACT osteoarthritis pain and functional impairment scales: ICOAP, KOOS-PS and HOOS-PS. *Clin Exp Rheumatol*. 2011;29(2):231-237.
22. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res*. 2002;52(2):69-77. doi:S0022399901002963 [pii]
23. Spinhoven P, Ormel J. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med*. 1997;27(2):363-370.
24. Van Damme S, Crombez G, Vlaeyen J, Goubert L, Van den Broeck A, Van Houdenhove B. De Pain Catastrophizing Scale: psychometrische karakteristieken en normering. 2000;3:211-222. *Gedragstherapie*. 2000;3:211-222.
25. Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess*. 1995;7(4):524-532. doi:10.1037/1040-3590.7.4.524
26. ten Klooster PM., Weekers AM., Eggelmeijer F., et al. Optimisme en / of pessimisme: factorstructuur van de Nederlandse Life Orientation Test Revised. *Psychol Gezondh*. 2010;38(2):89-100.
27. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - A practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):1-10. doi:10.1186/s12874-017-0442-1
28. Razmjou H, Finkelstein JA, Yee A, Holtby R, Vidmar M, Ford M. Relationship between preoperative patient characteristics and expectations in candidates for total knee arthroplasty. *Physiother Canada*. 2009;61(1):38-45. doi:10.3138/physio.61.1.38
29. Tolk JJ, Haanstra TM, Reijman M. Letter to the editor on "What so Scottish patients expect of their total knee arthroplasty?" *J Arthroplasty*. 2016;31(10):2374. doi:10.1016/j.arth.2016.05.023
30. Riddle DL, Golladay GJ, Hayes A, Ghomrawi HMK. Poor expectations of knee replacement benefit are associated with modifiable psychological factors and influence the decision to have surgery: A cross-sectional and longitudinal study of a community-based sample. *Knee*. 2017;24(2):354-361. doi:10.1016/j.knee.2016.11.009
31. Mahomed NN, Liang MH, Cook EF, et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *J Rheumatol*. 2002;29(6):1273-1279.
32. Lingard E a, Sledge CB, Learmonth ID. Patient expectations regarding total knee arthroplasty: differences among the United States, United kingdom, and Australia. *J Bone Jt Surg Am*. 2006;88-A(6):1201-1207. doi:10.2106/jbjs.e.00147
33. Hanusch BC, O'Connor DB, Ions P, Scott a., Gregg PJ. Effects of psychological distress and perceptions of illness on recovery from total knee replacement. *Bone Jt J*. 2014;96 B(2):210-

216. doi:10.1302/0301-620X.96B2
34. Vissers MM, Bussmann JB, Verhaar JAN, Busschbach JJ V, Bierma-Zeinstra SMA, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: A systematic review. *Semin Arthritis Rheum*. 2012;41(4):576-588. doi:10.1016/j.semarthrit.2011.07.003
35. Flood AB, Lorence DP, Ding J, McPherson K, Black NA. The role of expectations in patients' reports of post-operative outcomes and improvement following therapy. *Med Care*. 1993;31(11):1043-1056. doi:10.1097/00005650-199311000-00006
36. Colloca L, Miller FG. Harnessing the placebo effect: the need for translational research. *Philos Trans R Soc Lond B Biol Sci*. 2011;366(1572):1922-1930. doi:10.1098/rstb.2010.0399
37. Robertsson O, Dunbar MJ, Pehrsson T, Knutson K, Lidgren L. Patient satisfaction after knee arthroplasty: a report on 27,372 knees operated on between 1981 and 1995 in Sweden. *Acta Orthop Scand*. 2000;71(3):262-267.
38. Berliner JL, Brodke DJ, Chan V, SooHoo NF, Bozic KJ. Can Preoperative Patient-reported Outcome Measures Be Used to Predict Meaningful Improvement in Function After TKA? *Clin Orthop Relat Res*. 2016;475(1):1-9. doi:10.1007/s11999-016-4770-y
39. Dutch Arthroplasty Register (LROI). *LROI Annual Report 2018*. 's-Hertogenbosch, The Netherlands; 2018.

Abstract

Aims

The aim of this study was to assess the current available evidence about when patients might resume driving after elective, primary total hip (THA) or total knee arthroplasty (TKA) undertaken for osteoarthritis (OA).

Materials and Methods

In February 2016, EMBASE, MEDLINE, Web of Science, Scopus, Cochrane, PubMed Publisher, CINAHL, EBSCO and Google Scholar were searched for clinical studies reporting on 'THA', 'TKA', 'car driving', 'reaction time' and 'brake response time'. Two researchers (CAV and JJT) independently screened the titles and abstracts for eligibility and assessed the risk of bias. Both fixed and random effects were used to pool data and calculate mean differences (MD) and 95% confidence intervals (CI) between pre- and post-operative total brake response time (TBRT).

Results

A total of 19 studies were included. The assessment of the risk of bias showed that one study was at high risk, six studies at moderate risk and 12 studies at low risk. Meta-analysis of TBRT showed a MD decrease of 25.54 ms (95% CI -32.02 to 83.09) two weeks after right sided THA, and of 18.19 ms (95% CI -6.13 to 42.50) four weeks after a right-sided TKA, when compared with the pre-operative value.

Conclusion

The TBRT returned to baseline two weeks after a right-sided THA and four weeks after a right-sided TKA. These results may serve as guidelines for orthopaedic surgeons when advising patients when to resume driving. However, the advice should be individualised.

chapter 5

When is it safe to resume driving after total hip and total knee arthroplasty?

A meta-analysis of literature on post-operative brake reaction times.

C.A. van der Velden, J.J. Tolk, R.P.A. Janssen, M. Reijman

The Bone & Joint Journal. 2017 May;99-B(5):566-576.

Introduction

Total hip (THA) and total knee arthroplasty (TKA) for patients with osteoarthritis (OA) are among the most frequently performed surgical procedures in the western world.[1,2] It is estimated that the number of these procedures undertaken, for instance, in Holland will double during the next 15 years.[3]

In our increasingly demanding society, patients have high expectations and are eager to return to normal activities.[4] Return to driving is one of them, as it increases mobility and reduces social isolation and dependence on others.[5-8] Patient expectations are the most important predictor of post-operative satisfaction after THA and TKA.[9] Information about a realistic time frame for being able to drive again after the operation should be part of pre-operative counselling.

Dutch law notes that drivers must be able to control the vehicle, to avoid dangerous situations and react adequately if necessary.[10] In the United Kingdom, the Road Traffic Act states that drivers should not drive if their disability is likely to cause a danger to themselves or others.[11] It is the responsibility of the patient to judge when it is safe to drive again and they should consult a doctor or take a driving test if they are uncertain.[10,11] Dutch insurance companies are willing to insure patients when they start driving again if they have their doctors' permission.[10] According to a Dutch survey amongst orthopaedic surgeons, full weightbearing, proprioception and range of movement (ROM) should be taken into account when judging whether a patient can start driving again after fractures involving the lower limb.[12] An orthopaedic surgeon may be able to determine whether a limb can withstand the demands of driving, but other ergonomic, environmental and physical factors should be taken into account.[13] There should therefore be official guidelines.[12]

In order to drive safely, anyone must be able to make an emergency stop which requires an optimal reaction time and adequate grip on the steering wheel and power on the brake.[14] One measure of emergency braking is the total brake response time (TBRT), which records the total time required to press the brake pedal after a stimulus. It may be assessed post-operatively in different ways. A comparison of the pre- and post-operative times is frequently used.[5-7,15-25]

Current recommendations for driving after THA or TKA are largely based on older prospective studies. In these studies, post-operative TBRT was normalised to baseline within four to eight weeks postoperatively.[8,16,21,25] Despite the research in the last few years, there is still no consensus on when it is safe to drive after THA or TKA.[6,17-20,24]

The primary aim of the present study was to determine when the TBRT returns to pre-operative levels after THA and TKA. A secondary aim was to investigate the time that patients record resuming driving.

Materials and Methods

A systematic literature review and meta-analysis was performed, according to the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) guidelines.[26] Studies were eligible if they met the following criteria: adult patients with primary elective THA or TKA for OA on either the right or the left side. Prospective or retrospective studies with at least one of the following post-operative outcomes were eligible; TBRT, reaction time, brake force or time to resuming driving. The full text of the article and original pre- and post-operative data had to be available. Studies that included patients without a driving license were excluded. There were no restrictions in publication status, language or date of publication.
A literature search was conducted by an experienced librarian (WMB) in February 2016 of EMBASE, MEDLINE, Web of Science, Scopus, Cochrane, PubMed Publisher, CINAHL, EBSCO and Google Scholar. Search terms included synonyms of 'total hip arthroplasty', 'total knee arthroplasty', 'car driving', 'reaction time' and 'brake response time'. The complete search strategy for MEDLINE is shown in appendix 1. Similar search strategies were used for the other databases. Two independent researchers (CAV and JJT) screened titles and abstracts for eligibility. They reviewed the full text of eligible studies, and their references, to identify additional suitable studies. If they disagreed on the inclusion of a specific study, consensus was reached through discussion. A final decision by a third reviewer (RPAJ) was obtained if necessary. Two reviewers (CAV and MR) independently assessed the risk of bias, using a modified Cochrane Collaboration's tool for prognostic studies for assessing bias.[27] The tool contained questions to assess selection, information and confounding bias (Table 1). Low risk of bias was given when studies scored low risk on all three aspects, moderate risk when studies scored low risk on two aspects and high risk when studies scored low risk on one or no aspect. In all studies, consensus was reached through discussion and, if necessary, a third reviewer was consulted (RPAJ).

One researcher (CAV) extracted data from each study which was selected. A data collection form was used for the following items: authors, year of publication, study design, the characteristics of the patients, pre- and post-operative data, follow-up time and outcome. Where there were missing values such as mean, standard deviation, 95% confidence interval or standard error, the original authors were contacted. If they did not respond, the studies could not be pooled and therefore were not included in the meta-analysis.

Statistical analysis

The homogeneity of the studies was considered with regards to the study population, methodological quality, length of follow-up, outcome measures and determinants studied. The primary outcome measure was the mean difference between the pre- and post-operative TBRT. These data were used in the meta-analysis. The secondary outcome measure of the time to resume driving post-operatively as reported by patients was used for a qualitative analysis. The statistical software Review Manager 5.3 (Cochrane, London, United Kingdom) was used to combine and calculate each TBRT. Subgroup analyses of different post-operative times, hip and knee, and the affected side were performed. The consistency of results across the pooled studies was estimated with I^2 statistics. The I^2 values were interpreted as follows: 0% to 30%, no to low heterogeneity; 30% to 50%, moderate; 50% to 75%, substantial; > 75%, considerable.27 A p-value of < 0.05 indicated high importance of the I^2 value. The fixed

effects model was used when there was no significant statistical heterogeneity and the random effects model was used if there was statistical heterogeneity ($I2 > 50\%$, $p = 0.05$). Patients have returned to pre-operative levels when the post-operative TBRT is equal to, or shorter than, the pre-operative TBRT.

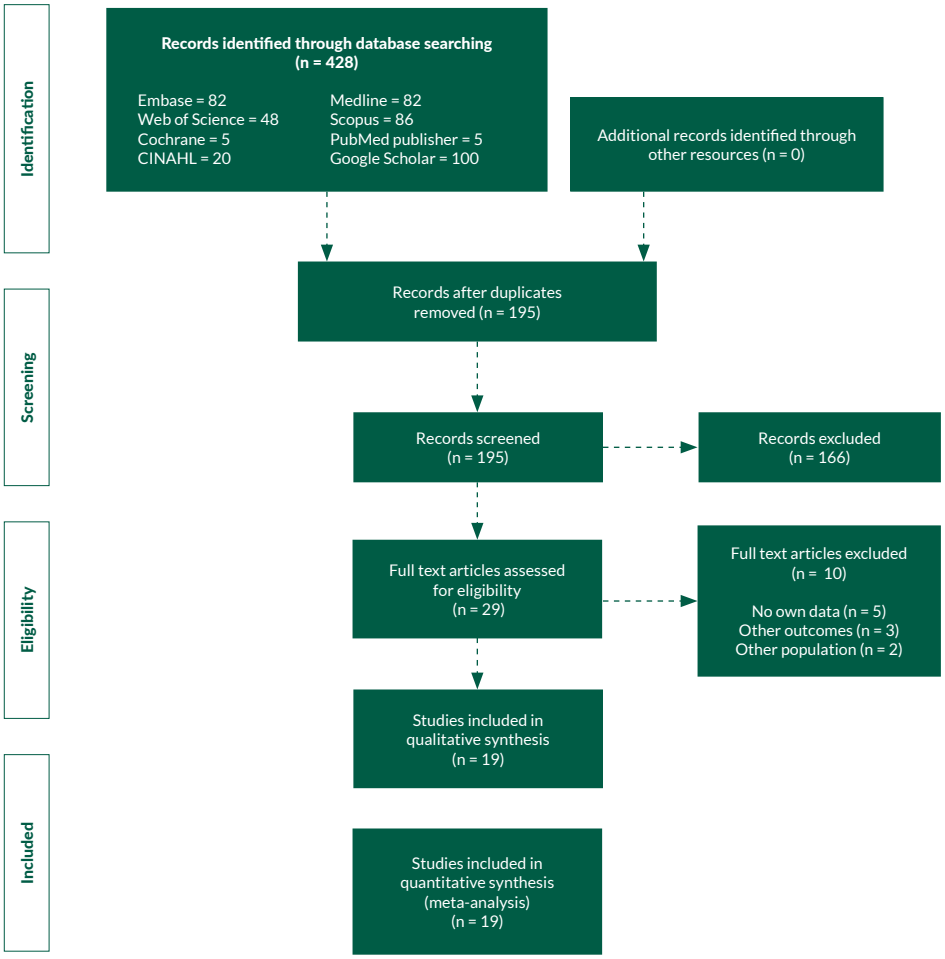


Figure 1. *Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart.*

Results

The PRISMA flowchart is presented in Figure 1. All included studies were prospective studies, published between 1988 and 2015. All patients used the right foot to press the brake pedal, except for in one study.[8] All patients were active drivers pre-operatively and were tested several times to limit learning effects. In studies of THA, patients took precautions (at every test point) in the seating position to avoid dislocation of the hip. Table 2 shows study characteristics.

The assessment of the risk of bias is summarised in Figure 2,[4-8,15-21,23-25,28-31] showing that one study had a high risk of bias,[4] six studies had a moderate risk[7,21,24,28-30] and 12 studies had a low risk[5,6,8,15,20,23,25,31].
The fixed effect model was used for statistical analysis. In one subgroup, two weeks after right-sided THA, the random effects model was used due to statistical heterogeneity. Table 3 presents the results of the TBRT in the studies and Figure 3 shows a forest plot of right-sided THA.[6,15,16,18,21,24,30] Two studies showed that the TBRT returned to the pre-operative level two weeks post-operatively.[6,24] After six weeks, the meta-analysis of these two studies showed a significant reduction in the post-operative TBRT compared with the pre-operative value.[16,18] Ganz et al. also measured the postoperative TBRT between four and six weeks.[16] To avoid overestimating the effect, the post-operative TBRTs were pooled in the six weeks subgroup. However, if their data were pooled in the subgroup of four weeks, the results already show a significant decrease in TBRT at this time.
Pooling of the results did not show a prolonged postoperative TBRT after left-sided THA (Fig. 4).[16,18,21] Data from one study showed a significant reduction in TBRT four weeks post-operatively.[16] The meta-analysis of two studies measuring TBRT eight to ten days after right-sided TKA showed a significantly longer TBRT, when compared with the pre-operative values.[7,31] Pooling of three studies showed that it returned to the pre-operative level four weeks post-operatively.[5,7,31] A forest plot shows right-sided TKA (Fig. 5).[5,7,17,19,20,31]

Pooling three studies after left-sided TKA did not show significant difference in the pre- and post-operative values (Fig. 6).[19,20,23]

Table 1. Modified Cochrane Collaboration's tool for assessing risk of bias

Bias type	
Selection bias **High risk** **Low risk**	No similar baseline measurements Similar baseline measurements
Confounding bias **High risk** **Low risk**	No report on side of surgery Side of surgery was given
Information bias **High risk ≤ 2 times yes** **Low risk ≥ 3 times yes**	Adequate follow up period? (>80% of patients returned to baseline TBRT at their last follow up measurement) Clearly defined outcome measures? Valid outcome measurements? (valid device or valid questionnaires to measure TBRT) Small loss to follow up? (< 20% of the patients)

Qualitative analysis

Qualitative analysis on patient-reported time to resume driving after TKA or THA**.** Of the 196 patients who underwent THA, 36 (18%) resumed driving within one month and 185 (94%) within three months. Abbas and Waheed reported that, of 130 patients, 105 (81%) resumed driving at six to eight weeks after THA (64% right-sided THA) and 127 (98%) within 12 weeks (62% right-sided THA).[28] A study by Ellanti et al showed that, of 98 patients after TKA (44% right-sided TKA), 77 (79%) resumed driving within six weeks and 95 (97%) within 12 weeks.[29] This is in accordance with the findings of Muh et al who found that, of 258 patients after right-sided TKA, 65 (25%) resumed driving within one month and 249 (97%) within three months.[4] Overall, the qualitative analysis showed that most patients resumed driving after THA or TKA within three months post-operatively (Table 4).

Table 4. Patient-reported time to resume driving after total hip or total knee arthroplasty

			Time to resume driving		
	Orthopaedic procedure	**Sample size (% right-sided)**	**< 4 wks (%)**	**< 6 to 8 wks (%)**	**< 12 wks (%)**
Abbas et al[28]	THA	130 (64)	-	105 (81)	127 (98)
Muh et al[4]	TKA	258 (100)	65 (25)	-	249 (97)
	THA	196 (100)	36 (18)	-	149 (94)
Ellanti et al[29]	TKA	98 (44)	-	77 (79)	95 (97)

THA, total hip arthroplasty; TKA, total knee arthroplasty

Discussion

This is the first meta-analysis on TBRT after elective primary, THA or TKA. The present study shows that TBRT returned to the pre-operative level two weeks after right sided THA[6,15,16,18,21,24,30] and four weeks after right-sided TKA[5,7,17,19,20,31]. Two high quality studies showed that it had significantly decreased six weeks after THA.[16,18] Previous systematic reviews reported a broader range for the resumption of driving based on TBRT: between two and eight weeks after right-sided THA, and between ten days and eight weeks after right-sided TKA.[14,22,32-35] This meta-analysis presents a more accurate estimation of time and suggests that patients may be able to drive earlier than previously recommended.

Data from those having a left-sided arthroplasty were also analysed, because TBRT might be affected by impaired function of the left knee on the performance of the contralateral braking leg.[36] These data showed that TBRT is minimally affected by a left-sided THA or TKA.[16,18-21,23] However, this evidence is not sufficiently robust to draw conclusions. Most patients reported resuming driving within three months after THA or TKA.[4,28,29] However, Muh et al found an increased risk of accidents in patients who resumed driving within three months.[4] No explanation was found for this result, and these results were not compared with those of a control group and no details on the seriousness of the accidents were given. Thus, these results are difficult to interpret.

The included studies have moderate to low risk of bias. Although most studies had similar patient characteristics such as gender, age, body mass index and years of driving, there was still possible heterogeneity between studies due to a time span of publication of 27 years, resulting in a variety of types of operation, implants and rehabilitation. There was no record of medication or other neurological or orthopaedic conditions that might affect the TBRT in some studies. The TBRT was measured at different post-operative times and a few studies had missing values.

This study has limitations. Although the TBRT is widely accepted as the best measure of competence when driving, some considerations should be kept in mind. First, although the relationship seems obvious, the exact correlation between TBRT and safe driving in real traffic is not known. Secondly, TBRT is divided into a reaction time and a movement time, and reaction can be neurologically impaired. It might have been better to measure movement separately, as it is dependent on pain, ROM and muscle strength. However, most studies did not assess this. Thirdly, different endpoints for the TBRT were used in the studies, varying from initial contact to a certain threshold of power in Newtons on the brake pedal. The time to reach a certain force might be lengthened by pain and weakness. Finally, the meta-analysis only included studies that compared pre- and postoperative TBRT. It is possible that patients with OA might already have a prolonged TBRT,[37] and pre-operative measurements might not be the best benchmark for comparison. A different approach would be to compare the post-operative TBRT with a set threshold. It has been shown that, when drivers are anticipating having to brake, the TBRT is about 700 ms to 750 ms, whereas in unexpected situations it is about 1250 ms.[37] Another approach would be to compare the post-operative TBRT with the TBRT of a control group. The TBRTs of studies which include a control group in this review ranged from 468 ms to 859 ms.[20,21,24,25] These large variations are due to different ways of measuring TBRT. Thus, there is no consensus on which fixed threshold should be used.

This review mainly focussed on the TBRT. However, there is little information about the correlation between TBRT and the actual safety of driving, the incidence of traffic accidents, brake force, strength, ROM and pain scores. Further studies are therefore required to objectify these outcomes in patients after THA and TKA. The length of the follow-up should be at least until the pre-operative values or those of a control group are reached. Furthermore, future research should also consider patients who have undergone revision arthroplasty.

Our findings suggest that patients seem to be able to make an emergency stop when using an automatic transmission two weeks after right-sided THA and four weeks after right-sided TKA. Further research would be required to determine when patients can resume driving after left sided THA and TKA. The results of this study can help surgeons advise patients on when patients might start driving after elective primary arthroplasty. However, driving is a complex activity requiring many skills, especially when travelling fast or using a manual transmission. Therefore, there are many variables that are involved in driving safely and the advice given to patients should be individualised. In order to indicate a safe time to resume driving doctors should also reflect on the individual characteristics of the patient including comorbidities, medication, the ability to weight-bear, the ROM, pain and muscle strength.

Take home message

• In this meta-analysis the TBRT returned to baseline at two weeks after right-sided THA and four weeks after right-sided TKA.
• This estimation is more accurate and suggests that patients may be able to drive earlier after THA and TKA than previously recommended.

Acknowledgements

We thank W. M. Bramer, biomedical information specialist from the Erasmus MC Rotterdam, for the literature search.

Table 2. Demographic characteristics of included studies in meta-analysis

Author/Design	Patient characteristics					Outcome measures			Follow-up period
	Patient group	n	Nfu	% Female	Mean age (SD)	Outcomes	Endpoint BRT	BRT comparison	
Dalury et al[5]/ Prospective	TKA Right	29	29	62%	66.0 (47-81)*	TBRT	Contact brake pedal	Pre-operative	8wk
Huang et al[17]/ Prospective	TKA Right	14	14	71%	63.1 (6.62)	TBRT	Fully depressed brake pedal until 50km/h	Pre-operative	4wk
Jordan et al[19]/ Prospective	TKA Right Left	45	40 20 20	50% 60%	69 (36-76)* 73 (55-86)*	TBRT	Contact with brake pedal	Pre-operative	1yr
Liebensteiner et al[20]/ Prospective	Control TKA Right Left	31 70 34 36	31 31 13 18	61% 44% 68%	52 (7.7) 65.9 (12.4) 65.7 (8.9)	TBRT	Contact with brake pedal	Pre-operative, Control group	8wk
Marques et al[23]/ Prospective	TKA Left	24	24	46%	63.2 (8.5)	TBRT	Contact with brake pedal	Pre-operative	10d
Marques et al[7]/ Prospective	TKA Right	21	21	57%	69.1 (7.8)	TBRT	Contact with brake pedal	Pre-operative	30d
Marques et al[31]/ Prospective	TKA Right	2	2	100%	68.5 (7.5)	TBRT	150 N brake pressure	Pre-operative	40d
Pierson et al[8]/ Prospective	TKA Bilateral Right Left	31 13 12 6	31 13 12 6	45%	68.6 (7.6)	TBRT	Fully depressed brake pedal	Pre-operative	9wk
Spalding et al[25]/ Prospective	Control TKA Right Left	14 18 12 6	14 18 12 6	Unknown	67.0 (52-85)* 74.0 (61-83)*	TBRT	100 N brake pressure	Pre-operative, Control group	10wk
Franz et al[15]/ Prospective	Control THA Right	22 14	22 14	Unknown	64.0 (5.8) 64.6 (4.7)	TBRT	500 N brake pressure	Pre-operative, Control group	3mth
Ganz et al[16]/ Prospective	THA Right Left	100	90 50 40	38% 52% 20%	63.3 (10.9) 65.1 (10.2)	TBRT	Contact with brake pedal	Pre-operative	1yr (26wk)
Hernandez et al[6]/ Prospective	THA Right	38	38†	Unknown	62.0 (10.5)	TBRT	Fully depressed brake pedal	Pre-operative	8wk
Jordan et al[18]/ Prospective	THA Right Left	47	40 20 20	50% 35%	67.7 (7.3) 65.5 (8.4)	TBRT	Contact with brake pedal	Pre-operative	1yr
MacDonald et al[21]/ Prospective	Control THA Bilateral Right Left	15 25	15 22 1 12 9	40% ? 33% 22%	? 61.0 (?) 58.0 (?)	TBRT	100 N brake pressure	Pre-operative, Control group	8wk (8mth)
Neumann et al[30]/ Prospective	Control THA Right	66 101	66 101	56% 39%	56.4 (7.6) 57.9 (7.8)	TBRT	Unknown	Control group	6wk
Ruel et al[24]/ Prospective	THA Right Wk 2 Wk 3 Wk4	90 30 29 31	90 30 29 31	70% 38% 52%	62.5 (7.1) 62.5 (7.8) 64.1 (8.6)	TBRT	Fully depressed brake pedal	Pre-operative	4wk
Abbas and Waheed[28]/ Prospective	THA Right Left	85 45	85 45	38%	Unknown	Questionnaire	-	-	12wk
Ellanti et al[29]/ Prospective	TKA Right Left	98 55 43	98 55 43	55%	59.5	Questionnaire	-	-	12wk
Muh et al[4]/ Prospective	TKA THA	258 196	258 196	72% 62%	70.1 (11.5) 70.5 (10.1)	Questionnaire	-	-	144mth

**range is given instead of standard deviation*
†of these 38 patients, 33 already reached baseline and were not included in analysis of the TBRT four weeks post-operatively
nFu, number of patients used in statistical analyses; SD, standard deviation; BRT, brake response time; TBRT, total brake response time; THA, total hip arthroplasty; TKA, total knee arthroplasty

Table 3. Summary of the results

Author	Group	Main results								
		Pre-operative	Postoperative TBRT							
			8-10 days	2wks	3wk	4wks	6wks	8wks	3mnd	1jr
Dalury et al[5]	TKA Right	530 (70)				490 (60)				
Huang et al[17]	TKA Right	1930 (1260-2520)*		2280 (1040-2660)*		1860 (1100-2190)*				
Jordan et al[19]	TKA Right Left	 587 (404-899)* 634 (488-868)*	 763 (443-1239)* 648 (486-933)*				600 (382-790)* 598 (477-899)*		555 (372-702)* 603 (457-826)*	522 (432-646)* 584 (446-715)*
Liebensteiner et al[20]	Control TKA Right Left	487 (116) 664 (64) 632 (46)		 674 (65) 642 (56)				 643 (55) 626 (50)		
Marques et al[23]	TKA Left	430 (82)	414 (63)							
Marques et al[7]	TKA Right	 420 (76)	 458 (87)			 428 (51)				
Marques et al[31]	TKA Right	 660 (121)	 786 (34)			 617 (37)	 657 (50)			
Pierson et al[8]	TKA				5,9% slower		12, 5% faster		17,5% faster	
Spalding et al[25]	Control TKA Right Left	710 (290) 720 (260)	 Increased Stable RT	 Increased Stable RT	 Increased Stable RT	 Increased Stable RT	 Increased Stable RT	 Preop RT Stable RT		
Franz et al[15]	THA Right	155					135		127	
Ganz et al[16]	 THA Right Left	 560 (120) 560 (140) 550 (90)	 590 (220) 630 (280) 530 (100)				 4-6wk 500 (90) 500(100) 500 (90)		 1 2jr: 450 (60) 450 (60) 460 (70)	 1jr: 480 (80) 470 (80) 480 (80)
Hernandez et al[6]	THA Right	635 (160)		576 (137)		501 (596)				
Jordan et al[18]	THA Right Left	 626 (296) 514 (194)	 777 (372) 549 (159)				 549 (129) 503 (95)		 517 (125) 483 (85)	 488 (101) 446 (70)
MacDonald et al[21]	Control THA Right Left	468 (136) 704 (135) 594 (135)						 656 (135) 495 (135)	8mnd: 591 (130) -	
Neumann et al[30]	Control THA Right	833 (184) 725 (114)			859 (212)		739 (142)			
Ruel et al[24]	THA Right	793 (180)		862 (258)	808 (254)	716 (118)				
Abbas and Waheed[28]	THA Right Left						Wk 6-8 51.5% 29.2%		 9.2% 7.7%	>12wk 0.8% 1.5%
Ellanti et al[29]	TKA						79%	12%	6%	3%
Muh et al[4]	 TKA THA					<4wk 25.19% 18.6%	1-3mth 71.32% 76.8%		<6mth 0.77% 3.09%	<1yr 0.39% 0%

**range is given instead of standard deviation (SD) due to missing data. The times are measured in milliseconds and given as mean and SD. TBRT, total brake response time; THA, total hip arthroplasty; TKA, total knee arthroplasty; RT, response time*

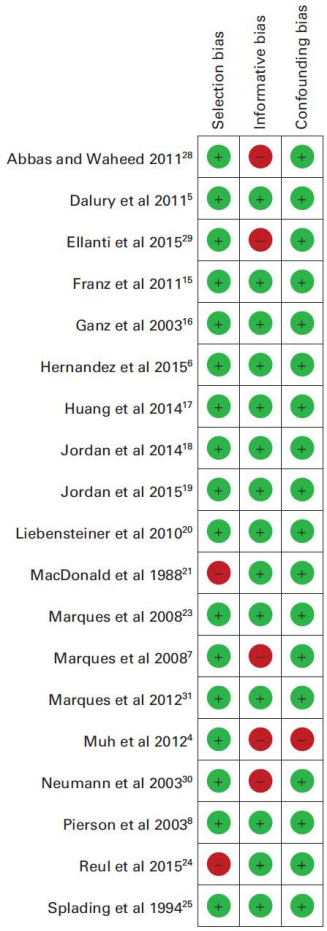


Figure 2. *Risk of bias assessment.*[4-8,15-21,23-25,28-31]

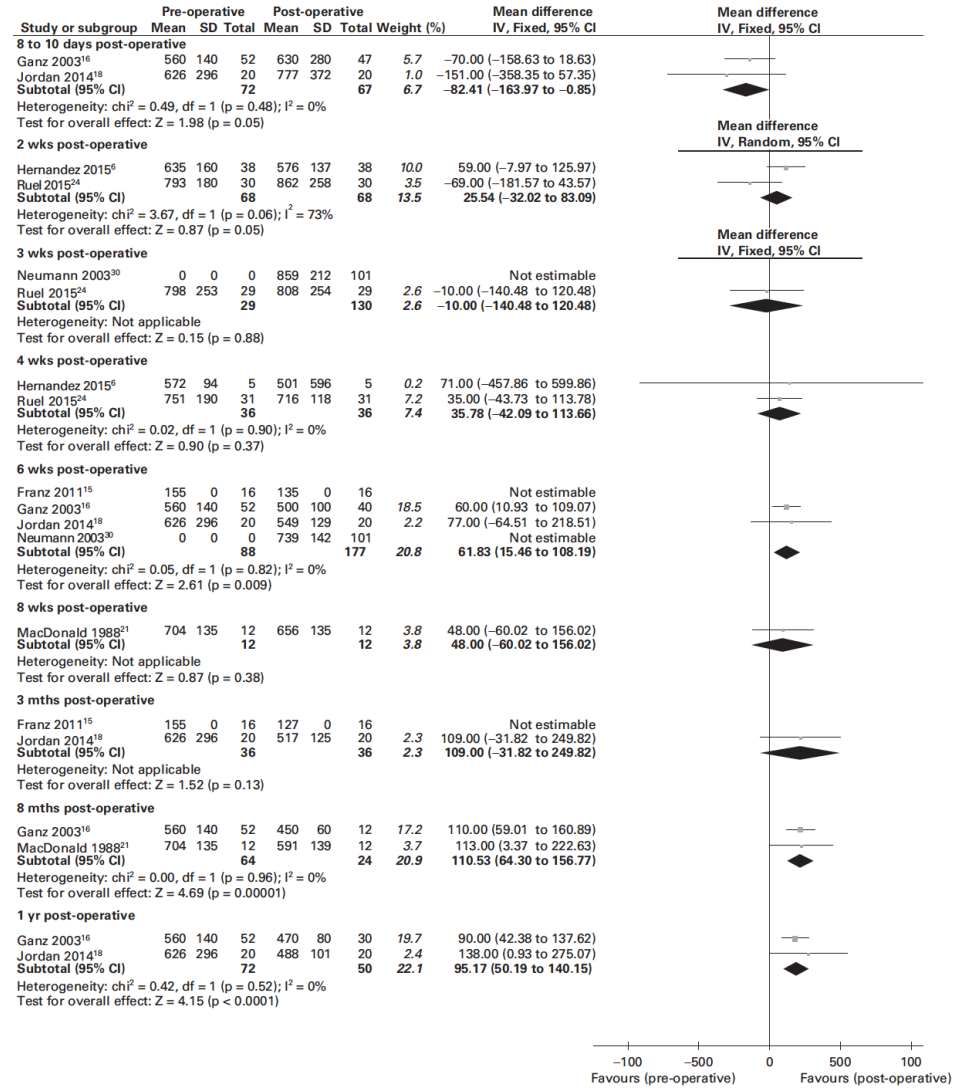


Figure 3. *Forest plot showing right-sided total hip arthroplasty (SD, standard deviation; IV, inverse-variance method; CI, confidence interval; df, degrees of freedom).*[6,15,16,18,21,24,30]

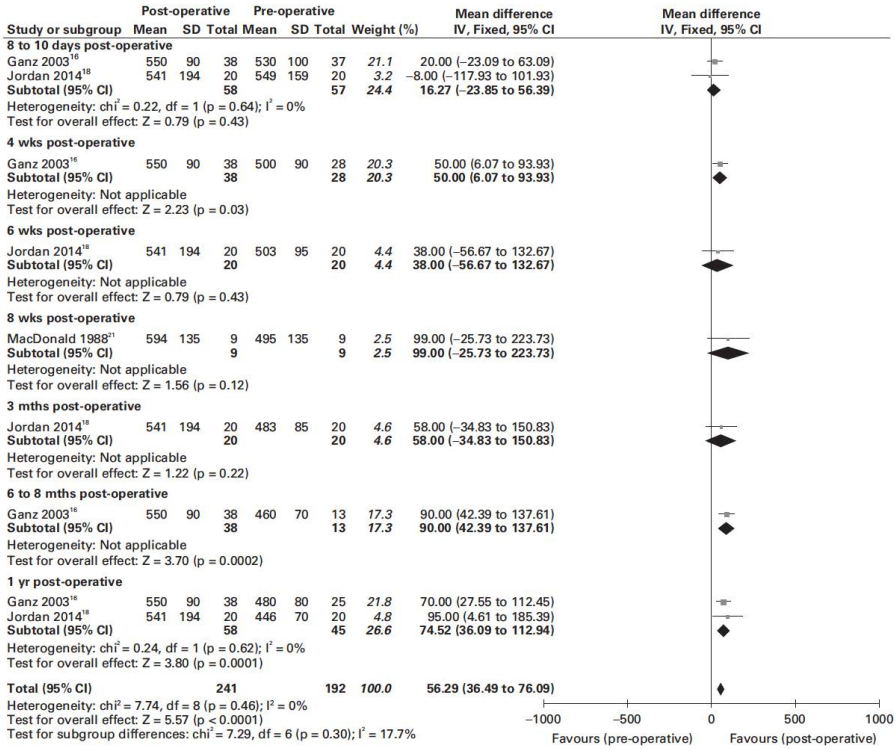


Figure 4. *Forest plot showing left-sided total hip arthroplasty (SD, standard deviation; IV, inverse-variance method; CI, confidence interval; df, degrees of freedom).*[16,18,21]

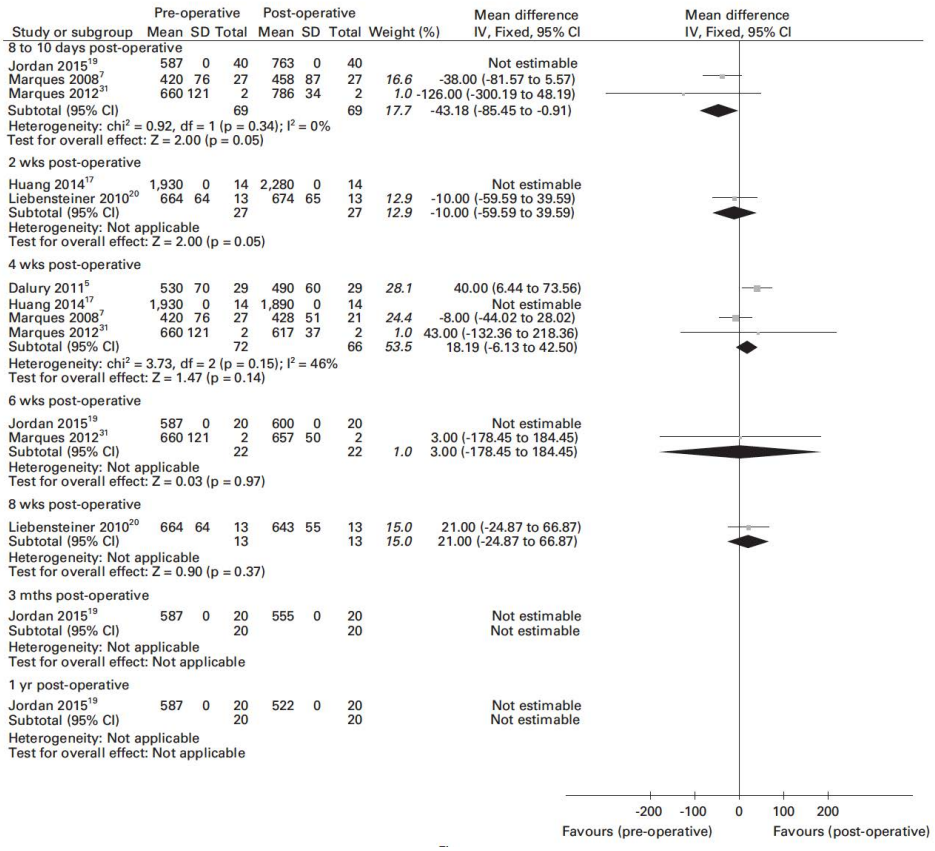


Figure 5. *Forest plot showing right-sided total knee arthroplasty (SD, standard deviation; IV, inverse-variance method; CI, confidence interval; df, degrees of freedom).5,7,17,19,20,31*

References

1. Nederlandse Orthopedische Vereniging. Richtlijn totale knieprothese. 2014.
2. Nederlandse Orthopedische Vereniging. Richtlijn Totale Knieprothese. 2010.
3. Otten R, van Roermund P, Picavet H. Trends in aantallen knie- en heuparthroplastieken: de vraag naar knie- en heupprotheses blijft voorlopig toenemen. *Ned Tijdschr Geneeskd*. 2010;154:A1534.
4. Muh S, Shishani Y, Streit J, et al. The Impact of Joint Replacement on Driver Function and Safety. *Open J Orthop*. 2012;02(03):121-125. doi:10.4236/ojo.2012.23022
5. Dalury DF, Tucker KK, Kelley TC. When can i drive?: Brake response times after contemporary total knee arthroplasty. In: *Clinical Orthopaedics and Related Research*. Vol 469. Springer New York LLC; 2011:82-86. doi:10.1007/s11999-010-1507-1
6. Hernandez VH, Ong A, Orozco F, Madden AM, Post Z. When is it safe for patients to drive after right total hip arthroplasty? *J Arthroplasty*. 2015;30(4):627-630. doi:10.1016/j.arth.2014.11.015
7. Marques CJ, Cabri J, Barreiros J, Carita AI, Friesecke C, Loehr JF. The Effects of Task Complexity on Brake Response Time Before and After Primary Right Total Knee Arthroplasty. *Arch Phys Med Rehabil*. 2008;89(5):851-855. doi:10.1016/j.apmr.2007.10.025
8. Pierson JL, Earles DR, Wood K. Brake response time after total knee arthroplasty: When is it safe for patients to drive? *J Arthroplasty*. 2003;18(7):840-843. doi:10.1016/S0883-5403(03)00326-7
9. Hamilton DF, Lane J V, Gaston P, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open*. 2013;3(4):1-8. doi:10.1136/bmjopen-2012-002525
10. Wegenverkeerswet 1994. http://wetten.overheid.nl/ BWBR0006622/2016-03-15. Accessed March 15, 2017.
11. *The Road Traffic Act 1988*. London: HMSO; 1988
12. Haverkamp D, Rossen NN, Maas AJJ, Olsman JG. Resuming driving after a fracture of the lower extremity: A survey among Dutch (orthopaedic) surgeons. *Injury*. 2005;36(11):1365-1370. doi:10.1016/j.injury.2005.05.039
13. Chen V, Chacko AT, Costello F V, Desrosiers N, Appleton P, Rodriguez EK. Driving after musculoskeletal injury. Addressing patient and surgeon concerns in an urban orthopaedic practice. *J Bone Joint Surg Am*. 2008;90(12):2791-2797. doi:10.2106/JBJS.H.00431
14. Fleury TR, Favrat B, Belaieff W, Hoffmeyer P. Resuming motor vehicle driving following orthopaedic surgery or limb trauma. *Swiss Med Wkly*. 2012;142. doi:10.4414/smw.2012.13716
15. Franz A, Mannel H, Brüggemann GP, Schmidt J. Bremsfähigkeit von Patienten nach endoprothetischer Versorgung des Hüftgelenks - Einfluss von Prothesendesign und präoperativem Status. *Z Orthop Unfall*. 2012;150(5):533-538. doi:10.1055/s-0032-1315245
16. Ganz SB, Levin AZ, Peterson MG, Ranawat CS. Improvement in Driving Reaction Time after Total Hip Arthroplasty. *Clin Orthop Relat Res*. 2003;(413):192-200. doi:10.1097/01.blo.0000072468.32680.ff
17. Huang HT, Liang JM, Hung WT, Chen YY, Guo LY, Wu WL. Timeframe for return to driving for patients with minimally invasive knee arthroplasty is associated with knee performance on functional tests. *BMC Musculoskelet Disord*. 2014;15(1). doi:10.1186/1471-2474-15-198
18. Jordan M, Hofmann UK, Grünwald J, et al. Influence of left- and right-side total hip arthroplasty on the ability to perform an emergency stop while driving a car. *Arch Phys Med Rehabil*. 2014;95(9):1702-1709. doi:10.1016/j.apmr.2014.03.009
19. Jordan M, Hofmann UK, Rondak I, Götze M, Kluba T, Ipach I. Brake response time is

significantly impaired after total knee arthroplasty: Investigation of performing an emergency stop while driving a car. *Am J Phys Med Rehabil*. 2015;94(9):665-676. doi:10.1097/PHM.0000000000000342

20. Liebensteiner MC, Kern M, Haid C, Kobel C, Niederseer D, Krismer M. Brake response time before and after total knee arthroplasty: A prospective cohort study. *BMC Musculoskelet Disord*. 2010;11. doi:10.1186/1471-2474-11-267
21. MacDonald W, Owen J. The Effect of Total Hip Replacement on Driving Reactions. *J Bone Jt Surg Br*. 1988;70(2):202-205.
22. Marecek GS, Schafer MF. Driving after orthopaedic surgery. *J Am Acad Orthop Surg*. 2013;21(11):696-706. doi:10.5435/JAAOS-21-11-696
23. Marques CJ, Barreiros J, Cabri J, Carita AI, Friesecke C, Loehr JF. Does the brake response time of the right leg change after left total knee arthroplasty? A prospective study. *Knee*. 2008;15(4):295-298. doi:10.1016/j.knee.2008.02.008
24. Ruel A V., Lee Y yu, Boles J, Boettner F, Su E, Westrich GH. A Novel Assessment of Braking Reaction Time Following THA Using a New Fully Interactive Driving Simulator. *HSS J*. 2015;11(2):143-147. doi:10.1007/s11420-015-9437-9
25. Spalding T, Kiss J, Kyberd P, Turner-Smith A, Simpson A. Driver Reaction Times After Total Knee Replacement. *J Bone Jt Surg Br*. 1994;76(5):754-756.
26. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol*. 2009;62(10):1006-1012. doi:10.1016/j.jclinepi.2009.06.005
27. Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*, *Cochrane Book Series*. West Sussex: Wiley-Blackwell; 2008
28. Abbas G, Waheed A. Resumption of car driving after total hip replacement. *J Orthop Surg (Hong Kong)*. 2011;19(1):54-56. doi:10.1177/230949901101900112
29. Ellanti P, Raval P, Harrington P. Return to driving after total knee arthroplasty. *Acta Orthop Traumatol Turc*. 2015;49(6):593-596. doi:10.3944/AOTT.2015.14.0307
30. Neumann U, Krause P, Erler K, Bruckner L. Change in reactivity in car drivers after artificial hip joint implantation. *Aktuelle Traumatol*. 2003;33:135-137.
31. Marques C, Santos R, Gamboa H, Lampe F. Brake response time before and after total knee arthroplasty: Tracking Possible Effects of the Surgery Technique on Motor Performance: Report of Two Cases [conference paper]. *Proc Int Conf Bio-inspired Syst Signal Process*. 2012:419–422.
32. Cooper JM. Clinical decision making: doctor, when can I drive? *Am J Orthop (Belle Mead NJ)*. 2007;36(2):78-80.
33. Goodwin D, Baecher N, Pitta M, Letzelter J, Marcel J, Argintar E. Driving after orthopedic surgery. *Orthopedics*. 2013;36(6):469-474. doi:10.3928/01477447-20130523-08
34. MacLeod K, Lingham A, Chatha H, et al. When can I return to driving?: A review of the current literature on returning to driving after lower limb injury or arthroplasty. *J Bone Jt Surg - Ser B*. 2013;95(3):290-294. doi:10.1302/0301-620X.95B3.29540
35. DiSilvestro KJ, Santoro AJ, Tjoumakaris FP, Levicoff EA, Freedman KB. When Can I Drive After Orthopaedic Surgery? A Systematic Review. *Clin Orthop Relat Res*. 2016;474(12):2557-2570. doi:10.1007/s11999-016-5007-9
36. Berth A, Urbach D, Awiszus F. Improvement of voluntary quadriceps muscle activation after total knee arthroplasty. *Arch Phys Med Rehabil*. 2002;83(10):1432-1436. doi:10.1053/apmr.2002.34829
37. Hofmann UK, Jordan M, Rondak I, Wolf P, Kluba T, Ipach I. Osteoarthritis of the knee or hip significantly impairs driving ability (cross-sectional survey). *BMC Musculoskelet Disord*. 2014;15(1). doi:10.1186/1471-2474-15-20

Abstract

Introduction

Unfulfilled preoperative expectations have a strong influence on the outcome after total knee arthroplasty (TKA). More insight into determinants of the level of expectations is useful in identifying patients at risk for having expectations of the treatment result that are too high or too low. This information can be used in optimizing pre-operative expectation management. The aim of the current study was to analyse to what extent pre-operative outcome expectations of TKA patients are affected by psychological factors, demographic factors, pain, physical function and general health status.

Methods

We performed a cross-sectional analysis of 204 patients with symptomatic and radiographic knee OA, scheduled for primary TKA. Outcome expectations were measured using the Hospital for Special Surgery knee replacement expectations survey. Independent variables included were age, sex, body mass index and patient reported outcome measures for pain, physical function, quality of life, anxiety, depression, catastrophizing, optimism and pessimism. Multiple linear regression analyses were used to evaluate associations between these variables and pre-operative outcome expectations.

Results

Female sex, higher age, higher depression score and duration of complaints > 50 months showed to be significant predictors of lower expectations for the treatment outcome after TKA. Baseline pain and function scores were not related to the level of pre-operative expectations.

Conclusion

The present study aids in identifying patients at risk for having either too high or too low expectations. This knowledge can be utilized in individualized expectation management interventions.

chapter 6

Outcome Expectations of Total Knee Arthroplasty Patients: The Influence of Demographic Factors, Pain, Personality Traits, Physical and Psychological Status.

J.J. Tolk, R.P.A. Janssen, T.M. Haanstra, M.C. van der Steen,
S.M.A. Bierma Zeinstra, M. Reijman

Journal of Knee Surgery. 2019 Jul 4. [Epub ahead of print]

Introduction

Symptomatic knee osteoarthritis (OA) can severely diminish quality of life and physical function.[1] Although, total knee arthroplasty (TKA) is generally considered an effective treatment option, still up to 20% of patients are not fully satisfied after surgery.[2,3] Unmet preoperative expectations have been reported as a strong factor influencing patient satisfaction after TKA.[2–4]

Several studies have shown that expectations can be modified by pre-operative education[5], but the best education strategy has not yet been identified[3,6]. In this education process patients with expectations of the treatment result that are too high or too low, have to be identified in order to be able to adjust their expectations accordingly.[7] If these patients can be specifically targeted, with an individualized education strategy, they are likely to benefit most from improved expectation management.

Previous research has not been able to consistently identify factors associated with pre-operative expectations of TKA patients. Expectations do not appear to be influenced by preoperative knee pain or patient-reported function scores. [8–10] It is suggested that psychological factors and personality traits may play significant roles in outcome expectations, but the available evidence on the effect of psychological factors on expectations of patients awaiting TKA is limited.[9]

The aim of the current study was to analyse the relationship between pre-operative factors and pre-operative outcome expectations in TKA patients. We hypothesized that psychological factors such as depressive symptoms, catastrophizing and optimism have a stronger relationship with the level of pre-operative expectations than demographic factors, pain, physical function and general health.

Methods

We performed a cross-sectional analysis of 204 patients scheduled for TKA between July 2016 and April 2018. Inclusion and exclusion criteria are described in Table 1. The present study is part of a trial on expectation management.[7] The [Blinded Manuscript] Medical Ethics Committee approved the study (registration code 15.108), and all patients signed an informed consent form.

Between July 2016 to April 2018 459 primary TKA were performed at [Blinded manuscript]. Of these patients, 204 could be included in the present study. Participation in the study was refused by 82 patients and 173 patients were excluded. Reasons for exclusion were contralateral TKA in 129 cases, indication for TKA other than OA in 11 cases, insufficient command of the Dutch language in 21 cases and planned staged or bilateral TKA in 12 cases. After enrolment, all patients completed a series of self-administered questionnaires. Demographic data, outcome expectations, health status, psychological status and personality traits were scored.

Demographic data

Patient demographics included age, sex, education level, body mass index (BMI), duration of complaints in months, status regarding employment for monetary reimbursement (yes/no), and co-morbidity scored using the Functional Comorbidity Index (FCI) [11]. Radiological OA severity was scored according to Kellgren & Lawrence (KL) grading system.[12] For use in the regression analysis the score was dichotomised into limited (KL grade 1 and 2) and evident radiological OA (KL grade 3 and 4).

Outcome expectations

We measured probability-based outcome expectations using the Dutch version of the Hospital for Special Surgery Knee Replacement Expectations Survey (HSS-KRES).[13] This 19-item self-administered survey measures probability-based outcome expectations in domains of pain, function, activities, and psychological well-being.[5,14] The expected improvement on each item is scored on a 5-point scale ranging from 0 (this expectation does not apply to me / I do not have this expectation) to 4 (complete improvement or return to normal). A total score can be calculated ranging from 0 to 100, with a higher score representing more positive expectations.[5,13] The survey is shown to be reliable and valid for the measurement of outcome expectations in TKA patients.[5,13–15]

Table 1. Inclusion and exclusion criteria

Inclusion criteria	Symptomatic and radiographic knee osteoarthritis indicated for a primary TKA
Exclusion criteria	Presence of TKA of the contralateral side Unicompartmental knee arthroplasty Staged or bilateral knee arthroplasty Insufficient command of the Dutch language Legally incompetent adults Presence of a medical illness that results in a life expectancy shorter than 1 year

Health status

The Dutch version of the Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form (KOOS-PS) is a patient-reported measure of physical function. The score consists of 7 questions scored a 5-point Likert. A normalized score can be calculated ranging from 0 indicating extreme symptoms to 100 indicating no symptoms. [16] The KOOS-PS has good reliability, validity and ability to detect change over time in knee OA patients.[16,17]

The Dutch version of the Oxford Knee Score (OKS) is a self-reported measure of pain and function. The questionnaire consists of 12 items on a 5-point Likert scale, of which the total score ranges from 12 to 60, with higher scores representing worse functional status. [18] The OKS is reproducible, valid and sensitive to clinically important changes.[17,18]

Generic health status is measured using the Dutch version of the EuroQol 5D-3L (EQ-5D).[19] The EQ-5D consists of 5 questions and a visual analogue scale (EQ-VAS). The questions are scored on a 3-point Likert scale, and a total score can be calculated. The lowest score indicates the worst health state possible and a score of 1 represents the best possible health state.[19] The EQ-VAS is scored from 0 ('Worst imaginable health state') to 100 ('Best imaginable health state'). The EQ-5D has good reliability and validity in knee OA patients.[20]

Pain during activity and rest is scored using an 11-point Numerical Rating Scale (NRS pain). Zero represents 'no pain' and a score of 10 represents 'worst imaginable pain'. The NRS pain has good reliability and responsiveness.[21]

Psychological status and personality traits

The Dutch version of the Hospital Anxiety and Depression Scale (HADS) is a valid and reliable measure of anxiety and depressive symptoms. [22,23] Seven 4-point scale questions that relate to anxiety and seven questions related to depression are rated. A 0 to 21 sum score for both subscales can be calculated, with 0 meaning no symptoms to 21 meaning severe symptoms. [22,23] The optimal cut-off score for the presence of both anxiety and depressive symptoms is ≥8. [22,23]

Catastrophizing is measured using the Dutch version of the Pain Catastrophizing Scale (PCS).[24] The PCS consists of 13 questions, with subscales for rumination, magnification and helplessness. Possible scores range from 0 (no catastrophizing) to 52 (extreme catastrophizing). The PCS is a reliable and valid self-reported measure of catastrophizing.[24,25]

The Dutch version of the Life Orientation Test - Revised (LOT-R) assesses optimism and pessimism.[26] This questionnaire has 10 items; three questions assess optimism, three pessimism, and the remaining four are filler items. Subscale scores can be calculated, and the total score is the result of adding the optimism to the inverted pessimism score. The LOT-R has satisfactory psychometric properties.[26]

Statistical Analysis

For descriptive statistics means with standard deviations (SD) for continuous variables were calculated and for discrete variables counts and percentages. To identify individual determinants of pre-operative outcome expectations we performed a multiple linear regression analysis. The dependent variable was the HSS-KRES score. As independent variables we included age, sex, education level, BMI, duration of complaints, co-morbidity (FCI), radiological OA, and the preoperative scores of KOOS-PS, OKS, NRS pain, EQ-5D, HADS, PCS and LOT-r. For potential predictors showing a non-linear relationship with the HSS-KRES score, the presence of a suitable cut-off value was explored. In the first step univariable regression analysis was performed with significance set at p=0.15. Potential predictors identified in the univariable analysis were entered into the multivariable model. The significance level in the multivariable analyses was set at a p-value of 0.05. Goodness-of-fit is reported using adjusted R2. Statistical analysis was performed with SPSS statistics version 24.0 (IBM corporation).

Results

All 204 patients under study completed the HSS knee replacement expectations survey. Patient characteristics, patient-reported function, pain score, general health score and measures of psychological status are reported in Table 2. Missing data accounted for less than 5% of cases for a limited number of independent variables (Table 2). Therefore we performed a complete case analysis.[27]

Table 2. Patient characteristics		
	Score	**Cases included in the analysis**
Age (years)	68.6 (9.3)	204
Sex, male [n (%)]	82 (40.2)	204
Side affected, right [n (%)]	106 (52)	204
BMI (kg/m^2)	29.0 (5.0)	204
Radiological OA severity [n (%)] KL 0 KL 1 KL 2 KL 3 KL 4 Evident radiological OA, KL 3 or 4 [n (%)]	 0 (0) 0 (0) 27 (13.2) 82 (40.2) 95 (46.6) 177 (86.8)	204
Duration of complaints Duration of complaints > 50 months [n (%)]	45.0 (67.1) 28 (13.7)	203
Education level [n (%)] Primary school Lower vocational education Pre-vocational secondary education Senior general secondary education Secondary vocational education Higher professional education University education	 12 (5.9) 67 (33.2) 38 (18.8) 8 (4.0) 38 (18.8) 32 (15.8) 7 (3.5)	202
Working status, yes [n (%)]	59 (28.9%)	204
FCI	2.5 (1.3)	204
Health status		
NRS pain At rest During activity	 4.9 (2.4) 7.9 (1.2)	204
EQ-5D Health scale Questions	 69.5 (20.7) 0.54 (0.28)	204
KOOS-PS	54.7 (12.9)	202
OKS	37.9 (6.3)	202
Psychological status and personality traits		
HADS Depression score Depression score ^{3} 8 [n (%)] Anxiety score Anxiety score ^{3} 8 [n (%)]	 4.1 (3.1) 29 (14.2) 4.2 (3.3) 27 (13.4)	202
PCS Rumination subscale Magnification subscale Helplessness subscale Total score	 6.6 (3.7) 1.9 (2.0) 6.6 (5.1) 15.1 (9.8)	197
LOTr Optimism subscale Pessimism subscale Total score	 9.2 (2.3) 4.1 (2.9) 17.1 (3.8)	201

Data are presented as mean and standard deviation between parentheses or reported otherwise as mentioned. BMI; Body Mass Index, KL; Kellgren and Lawrence, FCI; Functional Comorbidity Index, NRS; Numerical Rating Scale, KOOS-PS; Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form, OKS; Oxford Knee Score, HADS; Hospital Anxiety and Depression Scale, LOT-R; Life Orientation Test – Revised.

The mean overall survey score on the HSS-KRES was 70.9 (SD 17.9) with a range of 17.1 – 100.0. Distribution of expectation scores is shown in Figure 1. Highest expectations were scored for pain relief and improvement of the ability to walk of short and medium distances. Patients had the lowest expectations for improvement in kneeling, squatting, psychological well-being sexual activity and the ability to have paid work.

The univariate linear regression analysis showed 6 factors to be significant predictors of HSS-KRES score (Table 3). Duration of complaints could be included as a significant predictor using 50 months as cut-off value. In the multivariable analysis, 4 factors remained as shown in Table 3. Male sex, lower age, duration of complaints ≤50 months, and HADS depression score <8 were predictive of higher HSS-KRES scores. The model containing these 4 predictors had an adjusted R2 of 0.165.

Table 3. Regression analysis with HSS knee replacement expectation survey score.

	Univariable analysis				Multivariable analysis	
	R^2	B	95% CI	p-value	B	p-value
Age (years)	0.03	-0.3	(-0.60 – -0.08)	0.011	-0.3	0.023
Sex, female	0.03	-6.2	(-1.20 – -11.15)	0.015	-6.0	0.019
Duration of complaints >50 months	0.02	-7.2	(-14.3 – -0.1)	0.047	-9.8	0.006
Working status, yes	0.03	7.0	(1.59 – 12.35)	0.011	-	-
HADS Depression score ≥ 8	0.05	-12.1	(-19.00 – -5.17)	0.001	-10.4	0.003
LOTr Optimism	0.02	1.2	(0.09 – 2.24)	0.033	-	-

HSS; Hospital for Special Surgery, CI; confidence interval, HADS; Hospital Anxiety and Depression Scale, LOT-R; Life Orientation Test – Revised.
Note: Only predictors that were significant in the univariate analysis are shown.

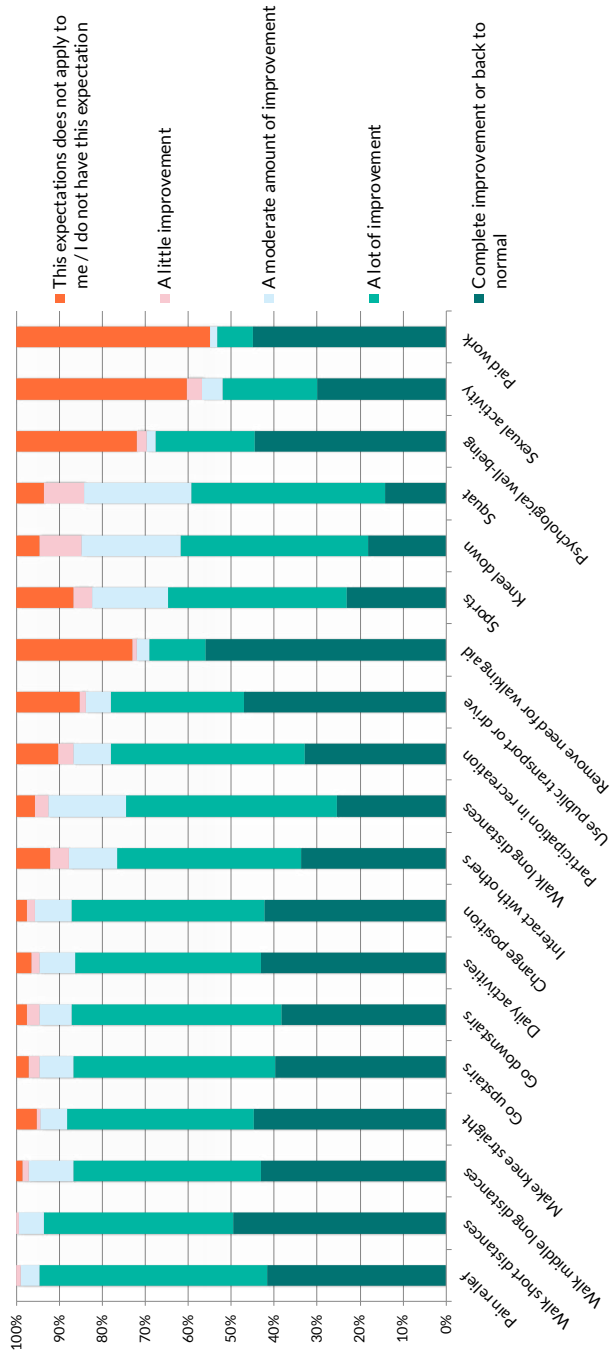


Figure 1. *Response to the Hospital for Special Surgery Knee Replacement Expectations Survey, questions are ordered by mean expectation score.*

Discussion

The most important finding of the present study is that female sex, higher age, HADS depression score ≥ 8 and duration of complaints >50 months are predictive for lower expectations for the treatment outcome after TKA. This only partly confirms our hypothesis that psychological factors are important predictors for level of expectations, because other items showed to be predictive as well and most psychological measures included in the analysis were shown not to be individual predictors of expectations.

Previous literature on the relationship between patient characteristics and expectations resulted in conflicting reports. This variation can for a large part be accounted for by the variation in patient cohorts and difference in definition and terminology.[8,28] Especially the distinction between value-based (what does a patient consider important); and probability-based expectations which we used for the present study (what does the patient think is the most likely result of treatment) has been the reason for confusion on this subject.[28,29] For studies reporting on predictors of probability-based outcome expectations, age and sex are reported as significant independent predictors of expectations. [9,28] Quite consistently preoperative patient-reported function scores are reported not to be related to the level of expectations.[9,30–32] Thus, the findings of the present study supports evidence from previous reports on determinants of probability-based outcome expectations.

In the present study, the HADS depression score was the only psychological factor that showed to be predictive of expectations. The finding that patients with more depressive symptoms have lower expectations is not surprising. These low expectations might be justified, as higher depression scores predict lower outcome after surgery.[33,34] On the other hand, the low expectations themselves might be partly responsible for the treatment outcome. Higher expectations are related to higher postoperative outcome,[8] and suggested explanations for this are anxiety reduction, better cooperation with treatment and beneficial coping mechanisms.[35,36] These positive effects are probably inversely related to the presence of depressive symptoms.[33,36] Therefore, the depressive symptoms, as well as the preoperative expectations, are a potential target for an intervention aimed at increasing the postoperative result in this group of patients. Future research should focus on developing an effective treatment strategy in this regard.

The present study showed that shorter duration of complaints significantly predicts higher outcome expectations, although the predictive value seems limited. To the authors' knowledge, this relation has not been described previously. It is known from previous work that patients with a shorter duration of complaints, have lower postoperative satisfaction scores.[2,37] As an explanation, Dunbar et al suggest that patients with a relatively short duration of complaints base their expectations of a TKA on their relatively high pre-diseased functional status. [2] Patients with long-standing complaints, might be more likely to accept a lower quality of health for themselves as this fits the frame of reference they have developed over time.[2] Expectation fulfilment, with subsequently a higher degree of postoperative satisfaction, can therefore be presumed more likely in the group of patients with longer duration of complaints.

The present study found no relation between preoperative knee pain or function scores and expectations. These findings reflect results from previous reports, where no relation of preoperative pain, [9,30,31] and function scores, [9,30–32] with the height of expectations is described. Given the known strong relationship between pre- and postoperative pain and function scores, these findings question the realism of patient expectations [9,38]. The previously posed suggestion that patients do not modulate their expectations on their personal functional situation and disease severity is supported by these findings. [9] This highlights the potential of improved expectation management to achieve more realistic expectations and subsequently a higher degree of expectation fulfilment.

A strength of the present study is the relatively large patient cohort and the wide range of potential predictors of patient expectations included in the analysis. Most previous studies identifying predictors for patient expectations are less concise, especially regarding psychological factors and duration of complaints.[8,9,28] Therefore the present study provides a clear overview of which factors do and do not play a role in influencing the level of patient expectations.
A possible limitation of the present study is that the study only identifies factors predicting higher or lower expectations for outcome after TKA. These patients might be at risk for having too high or too low expectations, but if this is actually the case cannot be directly drawn from the current study. Longitudinal follow-up is warranted to determine if these factors are related to expectation fulfilment and patient satisfaction. Furthermore, the current study was conducted in a Dutch setting and only osteoarthritis patients with an indication for primary TKA were included. It is known that patients from different countries have different expectations regarding TKA.[32] It is therefore possible that determinants of expectations show cross-cultural differences as well. This limits the generalizability of the results presented to some extent.
Not all patients that had a total knee replacement in our hospital during the study period could be included in the present analysis. The participation rate of eligible patient was 71.3%. This is an acceptable participation rate and baseline demographic and PROM scores did not show important difference to other populations of TKA patients.[39] Therefore, selection bias probably has not importantly influenced our conclusions. The results of the present study can be considered accurate for the group of patients scheduled for unilateral, primary TKA for symptomatic knee osteoarthritis. Generalizability to for example for revision TKA patients or patients with another indication than osteoarthritis, should be done with caution, as these patients might have a different pattern of expectations and determinants thereof.

Conclusions

Female sex, higher age, HADS depression score ≥ 8 and duration of complaints >50 months predict lower expectations for the treatment outcome after TKA. The present study aids in identifying patients at risk for having either too high or too low expectations. This knowledge can be utilized in individualized expectation management interventions.

References

1. Hoy D, Cross M, Smith E, et al. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis*. 2014;73(7):1323-1330. doi:10.1136/annrheumdis-2013-204763
2. 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):148-152. doi:10.1302/0301-620x.95b11.32767
3. 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. doi:10.1016/j.arth.2017.07.021
4. Noble PC, Conditt M a, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452(452):35-43. doi:10.1097/01.blo.0000238825.63648.1e
5. Mancuso C, Graziano S, Briskie L, et al. Randomized trials to modify patients' preoperative expectations of hip and knee arthroplasties. *Clin Orthop Relat Res*. 2008;466(2):424-431. doi:10.1007/s11999-007-0052-z
6. Tolk JJ, van der Steen M, Janssen RPA, Reijman M. Total knee arthroplasty: what to expect ? A survey of the members of the Dutch Knee Society on long-term recovery after total knee arthroplasty. *J Knee Surg*. 2017;30(6):612-616. doi:10.1055/s-0036-1593868
7. Tolk JJ, Janssen RPA, Haanstra TM, Bierma-zeinstra SMA, Reijman M. The EKSPECT study : the influence of Expectation modification in Knee arthroplasty on Satisfaction of PatiEnts : study protocol for a randomized Controlled Trial. *Trials*. 2018;19(437).
8. Jain D, Nguyen L-CL, Bendich I, et al. Higher patient expectations predict higher patient-reported outcomes, but not satisfaction, in total knee arthroplasty patients: a prospective multicenter study. *J Arthroplasty*. 2017;32(9S):S166-S170. doi:10.1016/j.arth.2017.01.008
9. Hepinstall MS, Rutledge JR, Bornstein LJ, Mazumdar M, Westrich GH. Factors that impact expectations before total knee arthroplasty. *J Arthroplasty*. 2011;26(6):870-876. doi:10.1016/j.arth.2010.09.010
10. Smith EJ, Soon V-L, Boyd A, McAllister J, Deakin A, Sarungi M. What do Scottish Patients Expect of Their Total Knee Arthroplasty? *J Arthroplasty*. 2016;31(4):786-792. doi:10.1016/j.arth.2015.10.021
11. Bjorgul K, Novicoff WM, Saleh KJ. Evaluating comorbidities in total hip and knee arthroplasty: available instruments. *J Orthop Traumatol*. 2010;11:203-209. doi:10.1007/s10195-010-0115-x
12. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16(4):494-502. doi:10.1136/ard.16.4.494
13. van den Akker-Scheek I, van Raay JJ a M, Reininga IHF, Bulstra SK, Zijlstra W, Stevens M. Reliability and concurrent validity of the Dutch hip and knee replacement expectations surveys. *BMC Musculoskelet Disord*. 2010;11(1):242. doi:10.1186/1471-2474-11-242
14. Mancuso CA, Sculco TP, Wickiewicz TL, Jones EC, Robbins L, Warren RF. Patients' expectations of knee surgery. *J Bone Jt Surg Am*. 2001;83-A(7):1005-1012.
15. Zywiel MG, Mahomed A, Gandhi R, Perruccio A V., Mahomed NN. Measuring expectations in orthopaedic surgery: A systematic review. *Clin Orthop Relat Res*. 2013;471(11):3446-3456. doi:10.1007/s11999-013-3013-8
16. de Groot IB, Favejee MM, Reijman M, Verhaar J a N, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. *Health Qual Life Outcomes*. 2008;6:16. doi:10.1186/1477-7525-6-16
17. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function:

International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Outcome. *Arthritis Care Res (Hoboken)*. 2011;63 Suppl 1(November):S208-28. doi:10.1002/acr.20632

18. Haverkamp D, Breugem SJM, 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 Orthop*. 2005;76(3):347-352. doi:10.1080/00016470510030814
19. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343.
20. Conner-Spady BL, Marshall DA, Bohm E, et al. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. *Qual Life Res*. 2015;24(7):1775-1784. doi:10.1007/s11136-014-0910-6
21. Ruyssen-Witrand A, Fernandez-Lopez CJ, Gossec L, Anract P, Courpied JP, Dougados M. Psychometric properties of the OARSI/OMERACT osteoarthritis pain and functional impairment scales: ICOAP, KOOS-PS and HOOS-PS. *Clin Exp Rheumatol*. 2011;29(2):231-237.
22. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res*. 2002;52(2):69-77. doi:S0022399901002963 [pii]
23. Spinhoven P, Ormel J. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med*. 1997;27(2):363-370.
24. Van Damme S, Crombez G, Vlaeyen J, Goubert L, Van den Broeck A, Van Houdenhove B. De Pain Catastrophizing Scale: psychometrische karakteristieken en normering. 2000;3:211-222. *Gedragstherapie*. 2000;3:211-222.
25. Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess*. 1995;7(4):524-532. doi:10.1037/1040-3590.7.4.524
26. ten Klooster PM., Weekers AM., Eggelmeijer F., et al. Optimisme en / of pessimisme: factorstructuur van de Nederlandse Life Orientation Test Revised. *Psychol Gezondh*. 2010;38(2):89-100.
27. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - A practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):1-10. doi:10.1186/s12874-017-0442-1
28. Razmjou H, Finkelstein JA, Yee A, Holtby R, Vidmar M, Ford M. Relationship between preoperative patient characteristics and expectations in candidates for total knee arthroplasty. *Physiother Canada*. 2009;61(1):38-45. doi:10.3138/physio.61.1.38
29. Tolk JJ, Haanstra TM, Reijman M. Letter to the editor on "What so Scottish patients expect of their total knee arthroplasty?" *J Arthroplasty*. 2016;31(10):2374. doi:10.1016/j.arth.2016.05.023
30. Riddle DL, Golladay GJ, Hayes A, Ghomrawi HMK. Poor expectations of knee replacement benefit are associated with modifiable psychological factors and influence the decision to have surgery: A cross-sectional and longitudinal study of a community-based sample. *Knee*. 2017;24(2):354-361. doi:10.1016/j.knee.2016.11.009
31. Mahomed NN, Liang MH, Cook EF, et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *J Rheumatol*. 2002;29(6):1273-1279.
32. Lingard E a, Sledge CB, Learmonth ID. Patient expectations regarding total knee arthroplasty: differences among the United States, United kingdom, and Australia. *J Bone Jt Surg Am*. 2006;88-A(6):1201-1207. doi:10.2106/jbjs.e.00147
33. Hanusch BC, O'Connor DB, Ions P, Scott a., Gregg PJ. Effects of psychological distress and perceptions of illness on recovery from total knee replacement. *Bone Jt J*. 2014;96 B(2):210-

216. doi:10.1302/0301-620X.96B2
34. Vissers MM, Bussmann JB, Verhaar JAN, Busschbach JJ V, Bierma-Zeinstra SMA, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: A systematic review. *Semin Arthritis Rheum*. 2012;41(4):576-588. doi:10.1016/j.semarthrit.2011.07.003
35. Flood AB, Lorence DP, Ding J, McPherson K, Black NA. The role of expectations in patients' reports of post-operative outcomes and improvement following therapy. *Med Care*. 1993;31(11):1043-1056. doi:10.1097/00005650-199311000-00006
36. Colloca L, Miller FG. Harnessing the placebo effect: the need for translational research. *Philos Trans R Soc Lond B Biol Sci*. 2011;366(1572):1922-1930. doi:10.1098/rstb.2010.0399
37. Robertsson O, Dunbar MJ, Pehrsson T, Knutson K, Lidgren L. Patient satisfaction after knee arthroplasty: a report on 27,372 knees operated on between 1981 and 1995 in Sweden. *Acta Orthop Scand*. 2000;71(3):262-267.
38. Berliner JL, Brodke DJ, Chan V, SooHoo NF, Bozic KJ. Can Preoperative Patient-reported Outcome Measures Be Used to Predict Meaningful Improvement in Function After TKA? *Clin Orthop Relat Res*. 2016;475(1):1-9. doi:10.1007/s11999-016-4770-y
39. Dutch Arthroplasty Register (LROI). *LROI Annual Report 2018*. 's-Hertogenbosch, The Netherlands; 2018.

Expectation management in clinical practice

part 03

Abstract

Background

One out of five patients is unsatisfied to some extent after total knee arthroplasty (TKA). Unmet expectations are the main driver of post-operative dissatisfaction. Improved pre-operative education on realistic expectations for long-term outcome after TKA, potentially leads to higher post-operative satisfaction. The effect of expectation modification on post-operative satisfaction in TKA patients has not yet been studied. The primary objective of the presented study is to examine whether an educational module on long-term recovery after TKA will improve patient satisfaction compared to usual pre-operative education.

Methods

The EKSPECT study is a randomized controlled trial. Patients with symptomatic and radiographic knee osteoarthritis indicated for a primary TKA will be randomized to usual pre-operative education (control group) or usual education plus an additional module on realistic expectations for long-term recovery (intervention group). Patients will be naïve to study objective and difference between study groups. Outcome expectations will be measured blinded for group allocation using the HSS Knee Replacement Expectations Survey at baseline (before the intervention), pre-operative (after the intervention) and fulfilment of expectations at 12-month follow-up. Baseline physical function, quality of life and psychological factors are measured using self-reported questionnaires. The primary outcome measure is satisfaction with treatment result at 12-month follow-up.

Discussion

The EKSPECT study will provide evidence on the effectiveness of an education module on long-term recovery after TKA, to improve treatment satisfaction. If beneficial, the education module is a simple intervention with a low burden for patients, which can easily be implemented in clinical practice.

Trial registration

Registered in the Dutch Trial Registry on March 17, 2016. Registration number: NTR5779. http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=5779.

chapter 7

The EKSPECT Study: The influence of Expectation modification in Knee arthroplasty on Satisfaction of PatiEnts: study protocol for a randomized Controlled Trial.

J.J. Tolk, R.P.A. Janssen, T.M. Haanstra, S.M.A. Bierma-Zeinstra, M.Reijman

Trials. 2018 Aug 14;19(1):437.

Background

Total knee arthroplasty (TKA) is a frequently performed procedure for patients with knee osteoarthritis. The Dutch arthroplasty registry reported an increase from 20.573 TKA in 2010, to 27.107 in 2016 in the Netherlands.[1] A further increase in these numbers is expected in the future due to aging of the Western population and growing number of people with overweight. [1]

TKA is considered an effective intervention for end-stage knee osteoarthritis. Considerable pain reduction, increase in physical function and quality of life can be achieved.[2,3] The treatment is relatively safe, cost-effective and excellent survival rates are reported, with prosthesis survival of more than 95% at 15 years follow-up.[2,4] Despite these favourable figures, the rate of satisfaction after TKA is consistently reported around 80%, leaving 1 in 5 patients unsatisfied to some extent after their knee surgery.[5,6]

Patients have multiple expectations regarding the outcome of TKA, mainly concerning relief of pain, improvement in physical functioning and improvement in psychosocial well-being.[7,8] Pre-operative expectations tend to be high, and are often overly optimistic.[9,10] Frequently a discrepancy exists between expectations of the patients and those of the surgeon, [11] and a substantial number of patients is reported to have unfulfilled expectations after TKA.[12] The fulfilment of pre-operative expectations on treatment outcome is reported as the main determinant of treatment satisfaction after TKA [5,6,13,14] Patients with unfulfilled expectations, are up to 10 times more likely to be dissatisfied with their treatment results.[13,14] These findings suggest that more realistic expectations potentially lead to higher post-operative satisfaction.

The current pre-operative patient education is predominantly focused on the process of care and the immediate post-operative period. Previous research has shown that structured pre-operative education on realistic expectations for long term recovery can alter pre-operative expectations.[15] After such an intervention, expectations of patients are reported to be lower and the rate of discordance between patients' and surgeons' expectations was reduced.[11,15] These findings suggest that specific education about post-operative outcome could lead to more realistic patient expectations. The effect of pre-operative expectation management on post-operative expectation fulfilment and ultimately better post-operative satisfaction after TKA has not yet been studied.

Trial objectives

The primary objective of this randomized controlled trial is to examine whether an additional education module on realistic expectations for long-term recovery of symptoms, physical functioning and psychological issues (intervention group) will improve patient satisfaction after TKA compared to usual pre-operative education (control group).

Furthermore, an analysis will be made to determine if the additional education module on pre-operative expectations of TKA patients leads to change in pre-operative outcome expectations and an increase in postoperative expectation fulfilment. An analysis will

be made on the relationship between expectation fulfilment and treatment satisfaction. Additionally, an explorative analysis will be performed on the effect of the additional education module in subgroups of patients, depending on age, gender, severity of symptoms, symptoms of depression, coping mechanisms and height of pre-operative expectations.

Methods

Study design

The influence of **E**xpectation modification in **K**nee arthroplasty on **S**atisfaction of **P**ati**E**nts, a randomized **C**ontrolled **T**rial (EKSPECT) study is a randomized clinical superiority trial, with a parallel group design and 1:1 allocation ratio. Patients will be randomized in group (a) usual education plus an additional module on realistic expectations for long-term recovery, or in group (b) usual education. Patients will be naïve to study objective and difference between study groups. Measurements will be performed blinded for group allocation at baseline, on day of admission and 12 months after TKA procedure. A flow chart of the study

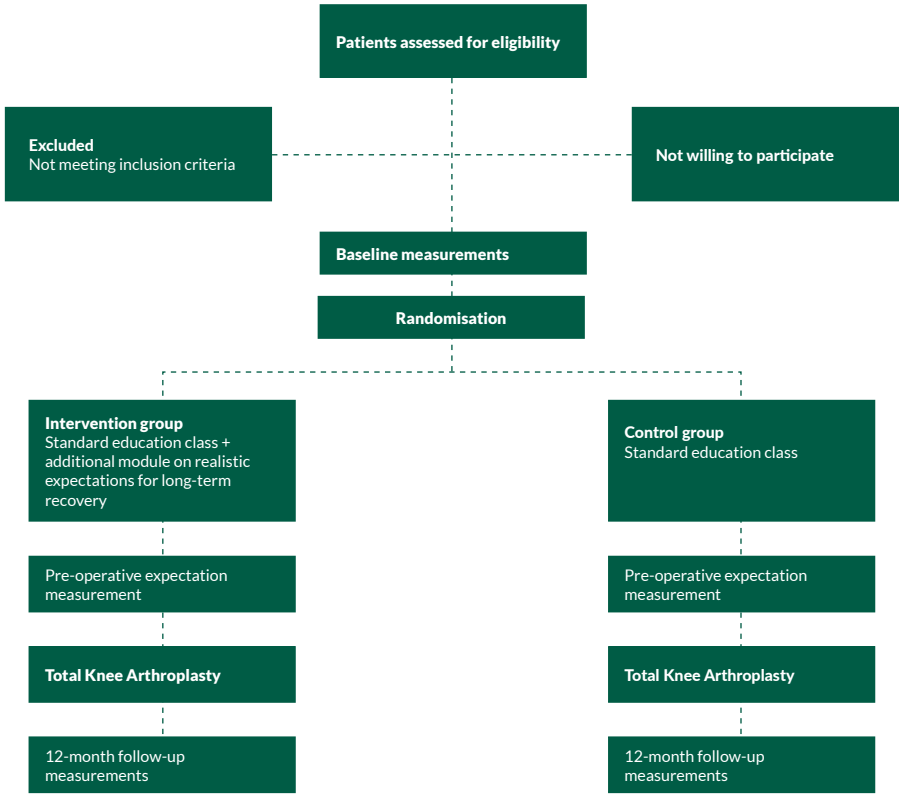


Figure 1. *EKSPECT study Consolidated Standards of Reporting Trials (CONSORT)-style flow diagram.*

procedures can be found in figure 1. *The* study has been reviewed and approved by the *Máxima Medical Centre* Medical Ethics Committee (registration code NL54671.015.15). The study has been registered in the Dutch Trial Registry (registration number NTR5779) and has not been amended. If protocol amendments are conducted, these will be updated on the Trial Registry record.

Setting

The study will be conducted at the department of orthopaedic surgery and trauma of the *Máxima Medical Centre*. This is a large non-academic teaching hospital where approximately 350 primary TKA are performed annually. In total 5 experienced orthopaedic surgeons perform these procedures.

Study population

Patients eligible for this trial are patients presenting at the outpatient clinic of department of orthopaedic surgery at the *Máxima Medical Centre*, with clinical and radiological knee osteoarthritis, indicated and scheduled for a TKA. The indication for a TKA is set according to the guideline of the Dutch Orthopaedic Society.[16] This guideline recommends considering TKA only in patients with radiological knee OA Kellgren and Lawrence $\geq$2 and pain and functional impairment with influence on quality of life, work and/or social life.[16]

Inclusion criteria

In order to be eligible to participate in this study, a subject must meet the following criteria:
- Symptomatic and radiographic knee osteoarthritis indicated for a primary TKA

Exclusion criteria

A potential subject who meets any of the following criteria will be excluded from participation in this study:

- Presence of a medical illness that results in a life expectancy shorter than 1 year,
- Presence of TKA of the contralateral side
- Unicompartmental knee arthroplasty
- Staged or bilateral knee arthroplasty
- Insufficient command of the Dutch language
- Legally incompetent adults

The exclusion criteria were chosen because the primary outcome will be measured at 1-year follow-up and sufficient proficiency of the Dutch language is necessary for understanding of the intervention under study and the completion of patient reported outcome measures. Furthermore, patients that already have a TKA on the contralateral side will be excluded because their personal experience with knee arthroplasty is known to result in considerable bias on the expectations of outcome after TKA.[17] Patients of whom at baseline it is already evident that they will undergo TKA during the study period (either bilateral or staged) are excluded. These patients are excluded because contralateral knee arthroplasty during the study period is likely to influence satisfaction and functional outcome scores of the index knee as well.

Recruitment

After being indicated for TKA by their orthopaedic surgeon and placed on the waiting list for surgery, eligible patients for the study will be screened by a research nurse. When in- and exclusion criteria are met, the patient will be asked to participate in the study. Written information on the study objective and procedures shall be provided, and patients will have the possibility to ask additional questions when necessary. The patient will be informed that two education modules are compared in this study, but the difference between the usual care and intervention module is not specified. During this process, patients are blinded for the hypothesis of the study to avoid bias and patient preference for one of the 2 modules. If a patient is willing to participate, an informed consent form will be signed and baseline measurements registered.

Randomization

The randomization procedure will be performed after the patient is cleared for surgery in pre-operative screening by the anaesthesiologist, and a definitive date for surgery has been determined. Allocation to type of intervention takes place by receiving the consecutive randomization number from the coordinating investigator. A computer-generated randomization list will be used (Research Manager version 5.30.0.6, Cloud9 Software, Deventer, the Netherlands). Patients will be included by each of the 5 orthopaedic surgeons performing and indicating patients for TKA. Because information on outcome after TKA provided by each orthopaedic surgeon during the consultation might be different on details, block randomization with variable size of blocks, stratified for orthopaedic surgeon is used. This accounts for potential information bias in this regard. All clinical and trial staff will be blinded for allocation during the study procedure. Patients were invited for a specific education class by mail to avoid cross-over, and only the secretary sending the invitations letters for the education sessions is unblinded. Attendance to the education class in the intervention and control arm is recorded blinded using the data management software. Data analysis will be obtained and analysed blinded for group allocation. Only after data analysis, unblinding will take place.

Intervention

Theoretical framework / rationale for using an education module to increase postoperative satisfaction

Patient expectations have been defined as "anticipations that given events are likely to occur during or as a result of medical care", and expectations regarding treatment result are defined as "outcome expectations". [18–21]. The present study addresses probalistic outcome expectations; what does the patient think is the most likely long-term result after TKA? [18–20]

Different hypotheses on the effect of expectation modification on treatment outcome and satisfaction have been suggested. In general positive expectation of recovery has shown to be associated with better health outcomes [22,23] and lower recovery expectations increased the risk of persistent activity limitations.[24] The strength of the relation depends on the clinical conditions and the method of expectation measurement used. [22] Expectations that are too low or negative can result in less motivation to obtain full benefit from the surgery, and increasing patient expectations is suggested to improve treatment outcome.[25,26] Suggested explanations for this positive effect are that higher expectations result in anxiety reduction, better cooperation with treatment and beneficial coping mechanisms.[27,28] There is evidence

that expectations are a mechanism by which placebos have their effects. [25] Utilization of this placebo effect as an intervention to obtain positive health-related effects has been shown in laboratory settings with predominantly healthy volunteers,[28] but these positive results have not been consistently reproduced in clinical research.[29] To the authors' knowledge there are no intervention studies available on the effect of increasing expectations to improve treatment outcome in TKA patients. So, it remains to be seen if this potential positive effect actually occurs and to what extent outcome improvement can be obtained.

On the other hand unrealistically high expectations can result in discouraged patients and non-adherence with recommendations post-operatively. Total knee arthroplasty patients often have overly optimistic pre-operative expectations, when compared to surgeons [11] and when compared to actual outcome.[9,30] Furthermore unmet pre-operative expectations are strongly related to post-operative dissatisfaction. [5,6] These findings are in line with the expectancy-disconfirmation theory, which states that satisfaction is a function of expectations, perceived performance, and disconfirmation of beliefs.[31] Therefore, to the authors' opinion an education module should not result in overly optimistic expectations, as these pose the risk of expectation disconfirmation with subsequent patient dissatisfaction.

In conclusion, expectation management in knee arthroplasty patients is thought to have effect on post-operative satisfaction and outcome through various pathways. Whereas on the one hand expectations should be high enough to fully benefit from the placebo effect, on the other hand unmet expectations can result in higher dissatisfaction rates after treatment. 'Optimistic realism' is suggested as the appropriate balance between preventing dissatisfaction and optimizing context effects.[32] In the authors view both aspects are important, and it is key to identify the patients with unrealistic expectations; either too high or too low and adjust these accordingly. Therefore, the aim of the proposed additional education module is to achieve realistic expectations on long-term recovery after TKA.

Intervention arm

The intervention under study is a joint-specific educational module on long-term recovery after TKA (12 months postoperative). The additional education module is an extension to the pre-operative education program as described for the control group. Information in the education module is based on literature study, expert opinion of TKA surgeons in our own clinic and a survey among members of the Dutch Knee Arthroplasty Society.[33] The final module was written and approved by an expert panel consisting of an experienced knee surgeon (RJ), orthopaedic surgery resident and PhD student (JT) and researcher specialized in osteoarthritis treatment outcome (MR).

Previous research has identified a set of important expectations that are often not fulfilled in TKA patients.[7,8,30] These items are incorporated in the Hospital for Special Surgery (HSS) Knee Replacement Expectations survey.[7,34] Items addressed in this survey were used as the framework for the education module. The module describes what patients can generally expect 12 months post-surgery, concerning the amount of pain, functioning in daily life (e.g. walking, chair rising, stair climbing); performing social activities (e.g. hobbies, sport activities); and psychological well-being (e.g., psychologic well-being, interactions with others). Information is provided for the most likely outcome for the whole population of

TKA patients. Additionally, modifying factors are addressed that predict higher or lower outcome for an individual patient; age, medical co-morbidity, body mass index (BMI), psychosocial factors, pain severity and pre-operative functional status.[35,36] The module consists of a group-based lecture of approximately 15 minutes, given by the senior author (JT, orthopaedic surgery resident and PhD student), and information on realistic outcome expectations after TKA in writing. When patients have additional questions or concerns regarding realistic expectations on treatment outcome, these will be addressed after the lecture. The module was formatted into the existing standard program with a session once a month for approximately five to ten patients, and is in concordance with recommendations on education class structure.[37]

In summary, the education module states the following. Overall good effect for long-term recovery after TKA can be expected.[33] Most improvement can be expected for the items pain relief, ability to perform daily activities and walking short to medium distances.[10,33,38] Significant pain relief is achieved in most patients after TKA, nevertheless, some residual pain is common.[10,38] Impairment in daily activities is likely to decrease. At 12 months much better or better activity of daily living function is reported by more than 90% of patients, only 4-8% of patients report some problems in daily living.[10,39] A large improvement or return to normal can be expected on walking short and medium distances.[10,33,40] For longer distances, some limitations are likely to remain.[10,33] Least improvement can be expected for the activities of kneeling, squatting, stair negotiation and the ability to exercise or participate in sports.[33,38] Approximately 80% of patients reports to be unable to kneel or squat without knee symptoms after TKA.[41,42] The rate of return to sport is dependent on patient characteristics and type of sports. The intensity of activity to which patients return tends to be less than before surgery; 94% is able to do low-impact sports but only 43% is reported to return to high-impact sports.[43,44]

The above summary is not exhaustive, but in the actual education module realistic expectations for all items of the HSS knee replacement survey are addressed. Improvement can be expected in domains of pain, function, activities, and psychological wellbeing but return to normal is not likely to occur. Limitations can predominantly be expected in more demanding physical activities and sports.

Control arm

The standard pre-operative education program consists of information about the perioperative period, but does not include information on long term recovery. The admission process, details about the anaesthesia process, surgical technique, complications, pain management, the direct postoperative recovery and rehabilitation during the first 6 weeks postoperative are addressed. During a 120-minute multidisciplinary pre-operative class an anaesthesiology assistant, physiotherapist, orthopaedic nurse and orthopaedic surgeon teach a 30-minute module each. The information is summarized in a brochure for the patients. This education module currently is the standard care in *Máxima Medical Centre* for all total knee arthroplasty patients.

Outcome measurement
Outcome measures used at the different timepoints can be found in figure 2.

Primary outcome measure
The primary outcome measure is a numeric rating scale (NRS) score for satisfaction with treatment result at 12-months follow-up. The NRS satisfaction score is a self-reported measure for patient satisfaction.[45] Patients are asked to answer the question 'How satisfied are you (in general) about the result of your knee operation?' on a scale ranging from 0 (very dissatisfied) to 10 (very satisfied).[45] 'Very satisfied with treatment result' is defined as an NRS satisfaction score of ≥8.

Patient characteristics
At baseline patient characteristics will be obtained: age, gender, side affected, height, weight, education level, duration of complaints and radiological osteoarthritis severity scored according to Kellgren & Lawrence grading system.[46]

Secondary outcome measures
Hospital for Special Surgery (HSS) Knee Replacement Expectations survey
Patient expectations will be evaluated by the Dutch version of the HSS Knee Replacement Survey.[34] Measurements will be obtained at baseline (after inclusion at the outpatient clinic, before pre-operative education), at admission for surgery (after the pre-operative education) and fulfilment of expectations 12 months post-operatively. The HSS Knee Replacement Expectations survey is a 19-item self-administered survey, measuring probability based outcome expectations in domains of pain, function, activities, and psychological wellbeing.[7,15] The survey is reliable, validated, and considered to be a high-quality expectation assessment instrument.[7,34,47] Patients will be asked how much improvement they expect for each item; the following response format will be used: 'complete improvement or back to normal', 'a lot of improvement', 'a moderate amount of improvement', 'a little improvement' or 'this expectation does not apply to me/I do not have this expectation'.[15,34] The total score ranges from 0 to 76, which will be recoded into a 100-point scale, with a higher score representing higher expectations. [15,34]
To assess to what extent expectations have been fulfilled 12 months postoperatively, the expectation fulfilment version of the HSS knee replacement expectations survey will be used.[12] Unfortunately no evaluation of the measurement properties of this modification are available. The perceived actual outcome is scored with the same answer options and score calculation as used for pre-operative measurement.[12]

Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form (KOOS-PS)
The *KOOS-PS* Dutch version is a 7-item knee-specific questionnaire for measurement of the construct physical function. From 5-point Likert scale questions, a normalized score is calculated ranging from no difficulty (0) to extreme difficulty (100).[48] KOOS-PS has good reliability, validity and ability to detect change over time in knee osteoarthritis patients.[48,49]

Oxford Knee Score (OKS)
The Dutch version of the OKS is a 12-item patient reported measure for assessment of pain and function after TKA. Each question consists of a 5-point Likert scale, leading to a total score ranging from a best functional score of 12 to the worst functional outcome score of 60.[50] It is short, reproducible, valid and sensitive to clinically important changes.[49,50]

EQ-5D

The Dutch version of the EuroQol 5D-3L (EQ-5D) is a self-reported questionnaire, measuring generic health status.[51] The EQ-5D comprises five questions scored on a 3 point Likert scale and a visual analogue scale (EQ VAS) where the endpoints are labelled 'Best imaginable health state' and 'Worst imaginable health state'. From the 5 questions a sum score can be calculated. 1 represents the best possible health state and lower scores imply a lower health state.[51] The EQ-5D has good reliability and validity in knee osteoarthritis patients.[52]

NRS pain

NRS score for pain during activity and at rest (NRS pain) during the past week will be measured on an 11-point scale. A score of 0 represents 'no pain' and a score of 10 represents 'worst imaginable pain'. The NRS has good reliability and responsiveness.[53]

Anchor question

At the follow-up 12 months postoperative a 7-point Likert scale anchor question will be scored for change in activities of daily living. The question 'how has your general daily functioning changed since the operation on your knee?' can be responded on a 7-point scale ranging from 1 (a lot worse) to 7 (very much improved).

Hospital Anxiety and Depression Scale (HADS)

The Dutch version of the HADS is a 14-item questionnaire for measurement of anxiety and depressive symptoms. [54,55] Seven items that relate to anxiety and seven items that relate to depression are rated on a 4-point scale. For both subscales, a sum score will be calculated ranging from 0 meaning no symptoms to 21 meaning severe symptoms.

The Pain Catastrophizing Scale (PCS)

The Dutch version of the PCS is a reliable and valid self-reported measure of catastrophizing.[56] Catastrophizing is defined as an exaggerated negative orientation toward noxious stimuli and is an important aspect of pain experience and coping.[56] The PCS consists of thirteen 5-point scale questions about thoughts and feelings on pain experience. Subscales for rumination, magnification and helplessness have been defined. A total score can be calculated ranging from 0 (no catastrophizing) to 52 (extreme catastrophizing).

Life Orientation Test-Revised (LOT-R)

The Dutch version LOT-R assesses the constructs dispositional optimism and pessimism and has satisfactory psychometric properties.[57,58] This self-reported measure consists of 10 items. Three items (1, 4, and 10) assess optimism, three items (3, 7, and 9) assess pessimism, and four are filler items. Response categories range from strongly agree to strongly disagree on a 5-point Likert scale. Scores for the 2 sub-scales can be calculated and the total score is calculated by adding the optimism and the inverted pessimism score.

Functional Co-morbidity Index (FCI)

Co-morbidity is scored using the FCI. The index consists of a list of 18 diagnoses associated with declining function.[59] One point is assigned to each diagnosis, and the points are summarized, giving the patient a score between 0 and 18.[59] The FCI is considered a reliable and valid tool for assessment of co-morbidity in osteoarthritis patients.[60]

Complications

At 1 year follow-up complications that have occurred will be scored as advised by the Knee Society.[61] This classification system allows structured reporting of occurrence and severity of 22 potential complications after TKA. [61] Adverse events are deemed unlikely due to the study design, occurrence will be monitored and reported.

Sample size and power calculations

The calculation of the number of patients needed for this trial, was based on the following assumptions. Specific preoperative information can lead to more realistic patient expectations, which subsequently leads to a higher probability of fulfilled expectations.[15] Furthermore, patients who have fulfilled expectations are more often satisfied with the results of TKA. [13,14] Therefore, we hypothesize that patients are more often (very) satisfied with treatment outcome one year postoperative. based on more realistic expectations after improved preoperative education on long-term recovery. Previous work has shown that 50% of patients is very satisfied after TKA, 23% somewhat satisfied, 11% neutral, 9% somewhat dissatisfied, and 7% was very dissatisfied with the outcome.[62] In our calculation, we used a power of 80% (type II error of 20%) and an alpha of 0.05 (type I error). To increase the proportion of very satisfied patients in the knee population with 20% (from 50% of very satisfied patients to 70% very satisfied patients), 90 patients are required in each group (180 TKA patients in total). The final total sample size required is 204 knee patients, to accommodate a 15% potential dropout rate.

Data management

All data will be handled confidentially and anonymized in compliance with the Dutch Personal Data Protection Act ('Wet Bescherming Persoonsgegevens'). Questionnaires are collected digitally and the patient study data will be stored coded using data management software (Research Manager version 5.30.0.6, Cloud9 Software, Deventer, the Netherlands). Each patient receives an anonymized study number that is used for all documentation, study reports and publications. The key of this study number will be handled by an independent researcher. All data will be stored during the study period, and when the study is finished, the research files are stored for 15 years in the *Máxima Medical Centre* research archive. The medical research ethics committee deemed the study as a 'low risk' study. Therefore, no data monitoring committee was recommended at the time of ethical approval. The medical research ethics committee will be informed yearly on the inclusion rate, adverse events and study results.

Statistical analysis

The primary analysis will be performed according to intention-to-treat principles following the original group allocation regardless of the intervention they did actually receive. Missing data will be accounted for by a multiple imputation technique. A secondary per-protocol analysis will be performed, including only those patients who completed the treatment originally allocated. Reasons for cross-over will be explored. Distribution of all variables will be tested by the Shapiro-Wilk test. For the normally distributed variables parametric tests will be used. For the variables that were not normally distributed non-parametric tests will be used.

Primary outcome will be analysed by using logistic regression analyses (patient satisfaction

with the 12 months results of TKA as dependent and intervention as independent variable). Variables of which a priori is known that they are associated with patient satisfaction, based on previous studies or based on a strong clinical rationale, will be considered as covariates in the primary analysis. These covariates are age, gender, BMI and baseline expectations as assessed by the HSS Knee Replacement Survey, anxiety, depression. The assumptions of constant variance and linear relationships will be assessed using scatter plots. Should any of these assumptions seriously fail, then transformation of the dependent or independent variable(s) (where applicable) will be used. The choice of which transformation (e.g. square root, logarithm) will be used based on the specific distribution of the residuals. Similar analyses will be performed for the secondary outcome parameters. Change in outcome expectations before and after pre-operative education (baseline vs pre-operative), and difference between intervention and control group in change in baseline and pre-operative expectations will be analysed using linear regression analyses. Statistical calculations will be made using IBM® SPPS software, version 19.0.

Discussion

The EKSPECT study is the first trial to analyse the potential of expectation management to increase post-operative satisfaction in TKA patients. Meeting pre-operative expectations is known to be of major influence on post-operative satisfaction after TKA. More realistic expectations will potentially lead to higher post-operative satisfaction. The present study will analyse the effect of improved pre-operative education, with specific attention to realistic expectations for long-term functional recovery after TKA, on post-operative satisfaction.

Strengths of the study include the sound methodological framework, with double-blind, randomized allocation and assessment of the intervention. Secondly, in previous research on expectation management poor measurement methods and inconsistent definition of constructs under study are common.[18,19] It is therefore recommended to clearly define the construct measured and the theoretical framework it derives from, to allow accurate interpretation of the results.[19] To the authors opinion these factors are adequately addressed in the present study design. Thirdly, the study procedures will be fully integrated in the current clinical process. This aids to the generalizability of the study results, limits the burden for the study population and increases the likelihood for patients to be willing to participate.

A limitation of the proposed study is that the content of the intervention education module is based on what patients undergoing TKA consider most important expectations,[7,8,30] and patient experiences after TKA as previously reported in literature. Although a patient-centred approach was used in the design, patients were not directly involved in the construction of the module.

Furthermore, a limitation of the proposed study could be that the intervention provides education on realistic expectations for the general TKA population. Potential individual modifiers influencing outcome after TKA are provided, but the prediction is not individualized. Currently available prediction models do not seem suitable for

individualized expectation management.[63] The existing outcome prediction tools mainly focus on identifying patients most likely not to benefit from TKA, but do not provide specific information on pain and functional outcome to guide pre-operative expectation management.[63] Furthermore, such an individual approach would increase the burden for patients and medical staff to a much larger extent than the proposed additional education module. Therefore, this pragmatic design was chosen to keep the burden for the patient as low as possible and increase the possibility of future implementation when positive results are found.

For the study osteoarthritis patients with an indication for primary TKA will included in a large non-academic hospital in the Netherlands. It is known that patients from different countries have different expectations regarding total knee arthroplasty. [64] Expectations and satisfaction rates are known to differ across indications for TKA as well.[6] These factors might limit the generalizability of the study results to some extent.

	STUDY PERIOD				
	Enrolment	**Allocation**	**Post-allocation**		**Close-out**
TIMEPOINT	Baseline, at inclusion	Randomisation	Pre-operative education	Pre-operative, post-education	12-months follow-up
ENROLMENT:					
Eligibility screen	X				
Informed consent	X				
Allocation		X			
INTERVENTIONS:					
Intervention group: *Additional educational module on long-term recovery*			X		
Control group: *Standard education program*			X		
ASSESSMENTS:					
Demographic; *age, gender, side affected, BMI, education level, duration of complaints, radiological osteoarthritis*	X				
NRS satisfaction					X
HSS knee replacement expectations survey	X			X	X•
KOOS-PS	X				X
OKS	X				X
EQ-5D	X				X
NRS pain	X				X
Anchor question					X
HADS	X				
PCS	X				
LOT-R	X				
FCI	X				
Complications					X

BMI; Body Mass Index. NRS satisfaction; Numerical Rating Scale for satisfaction with treatment result. HSS knee replacement expectations survey; Hospital for Special Surgery Knee Replacement Expectations survey KOOS-PS; Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form, OKS; Oxford Knee Score, NRS pain; Numerical Rating Scale for pain, HADS; Hospital Anxiety and Depression Scale. PCS; The Pain Catastrophizing Scale. LOT-R; Life Orientation Test-Revised, FCI; Functional Co-morbidity Index. * HSS knee replacement expectation survey at 12-month follow-up adapted for expectation fulfilment.

Figure 2. *Standard protocol items: recommendation for interventional trials (SPIRIT) summary of study timing and activities*

References

1. Dutch Arthroplasty Register (LROI). *Online LROI Annual Report 2017: 10 Years of Registration, a Wealth of Information.* 's-Hertogenbosch, The Netherlands; 2017.
2. Ruiz D, Koenig L, Dall TM, et al. The Direct and Indirect Costs to Society of Treatment for End-Stage Knee Osteoarthritis. *J Bone Jt Surg Am*. 2013;95(16):1473-1480. doi:10.2106/JBJS.L.01488
3. Tolk JJ, Janssen RPA, Prinsen CAC, et al. The OARSI core set of performance-based measures for knee osteoarthritis is reliable but not valid and responsive. *Knee Surgery, Sport Traumatol Arthrosc*. 2019;27(9):2898-2909. doi:10.1007/s00167-017-4789-y
4. McCalden RW, Hart GP, MacDonald SJ, Naudie DD, Howard JH, Bourne RB. Clinical Results and Survivorship of the GENESIS II Total Knee Arthroplasty at a Minimum of 15 Years. *J Arthroplasty*. 2017;32(7):2161-2166. doi:10.1016/j.arth.2017.02.006
5. Noble PC, Conditt M a, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452(452):35-43. doi:10.1097/01.blo.0000238825.63648.1e
6. 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):148-152. doi:10.1302/0301-620x.95b11.32767
7. Mancuso CA, Sculco TP, Wickiewicz TL, Jones EC, Robbins L, Warren RF. Patients' expectations of knee surgery. *J Bone Jt Surg Am*. 2001;83-A(7):1005-1012.
8. Smith EJ, Soon V-L, Boyd A, McAllister J, Deakin A, Sarungi M. What do Scottish Patients Expect of Their Total Knee Arthroplasty? *J Arthroplasty*. 2016;31(4):786-792. doi:10.1016/j.arth.2015.10.021
9. Mannion AF, Kämpfen S, Munzinger U, Kramers-de Quervain I. The role of patient expectations in predicting outcome after total knee arthroplasty. *Arthritis Res Ther*. 2009;11(5):R139. doi:10.1186/ar2811
10. 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 Orthop*. 2009;80(1):55-61. doi:10.1080/17453670902805007
11. Ghomrawi HMK, Mancuso C a., Westrich GH, Marx RG, Mushlin AI. Discordance in TKA expectations between patients and surgeons knee. *Clin Orthop Relat Res*. 2013;471(1):175-180. doi:10.1007/s11999-012-2484-3
12. Tilbury C, Haanstra TM, Leichtenberg CS, et al. Unfulfilled Expectations After Total Hip and Knee Arthroplasty Surgery: There Is a Need for Better Preoperative Patient Information and Education. *J Arthroplasty*. 2016;31(10):2139-2145. doi:10.1016/j.arth.2016.02.061
13. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: Who is satisfied and who is not? *Clin Orthop Relat Res*. 2010;468(1):57-63. doi:10.1007/s11999-009-1119-9
14. Hamilton DF, Lane J V, Gaston P, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open*. 2013;3(4):1-8. doi:10.1136/bmjopen-2012-002525
15. Mancuso C, Graziano S, Briskie L, et al. Randomized trials to modify patients' preoperative expectations of hip and knee arthroplasties. *Clin Orthop Relat Res*. 2008;466(2):424-431. doi:10.1007/s11999-007-0052-z
16. Nederlandse Orthopaedische Vereniging. Richtlijn totale knieprothese. 2014.
17. Hepinstall MS, Rutledge JR, Bornstein LJ, Mazumdar M, Westrich GH. Factors that impact expectations before total knee arthroplasty. *J Arthroplasty*. 2011;26(6):870-876. doi:10.1016/j.arth.2010.09.010

18. Haanstra TM, van den Berg T, Ostelo RW, et al. Systematic review: do patient expectations influence treatment outcomes in total knee and total hip arthroplasty? *Health Qual Life Outcomes*. 2012;10(1):152. doi:10.1186/1477-7525-10-152
19. Tolk JJ, Haanstra TM, Reijman M. Letter to the editor on "What so Scottish patients expect of their total knee arthroplasty?" *J Arthroplasty*. 2016;31(10):2374. doi:10.1016/j.arth.2016.05.023
20. Kravitz RL. Patients' expectations for medical care: an expanded formulation based on review of the literature. *Med Care Res Rev*. 1996;53(1):3-27.
21. Uhlmann RF, Inui TS, Carter WB. Patient requests and expectations. Definitions and clinical applications. *Med Care*. 1984;22(7):681-685. doi:10.1097/00005650-198407000-00011
22. Mondloch M V, Cole DC, Frank JW. Does how you do depend on how you think you'll do? A systematic review of the evidence for a relation between patients' recovery expectations and health outcomes. *Can Med Assoc J*. 2001;165(2):174-179.
23. Jain D, Nguyen L-CL, Bendich I, et al. Higher patient expectations predict higher patient-reported outcomes, but not satisfaction, in total knee arthroplasty patients: a prospective multicenter study. *J Arthroplasty*. 2017;32(9S):S166-S170. doi:10.1016/j.arth.2017.01.008
24. Iles RA, Davidson M, Taylor NF, O'Halloran P. Systematic Review of the Ability of Recovery Expectations to Predict Outcomes in Non-Chronic Non-Specific Low Back Pain. *J Occup Rehabil*. 2009;19(1):25-40. doi:10.1007/s10926-008-9161-0
25. Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H. The role of expectancies in the placebo effect and their use in the delivery of health care: A systematic review. *Health Technol Assess (Rockv)*. 1999;3(3):1-96. doi:10.3310/hta3030
26. Ebrahim S, Malachowski C, Kamal el Din M, et al. Measures of Patients' Expectations About Recovery: A Systematic Review. *J Occup Rehabil*. 2014;25(1):240-255. doi:10.1007/s10926-014-9535-4
27. Flood AB, Lorence DP, Ding J, McPherson K, Black NA. The role of expectations in patients' reports of post-operative outcomes and improvement following therapy. *Med Care*. 1993;31(11):1043-1056. doi:10.1097/00005650-199311000-00006
28. Colloca L, Miller FG. Harnessing the placebo effect: the need for translational research. *Philos Trans R Soc Lond B Biol Sci*. 2011;366(1572):1922-1930. doi:10.1098/rstb.2010.0399
29. Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. Influence of context effects on health outcomes: a systematic review. *Lancet*. 2001;357:757-762. doi:10.1016/S0140-6736(00)04169-6
30. Tilbury C, Haanstra TM, Leichtenberg CS, et al. Unfulfilled Expectations After Total Hip and Knee Arthroplasty Surgery: There Is a Need for Better Preoperative Patient Information and Education. *J Arthroplasty*. 2016;31(10):2139-2145. doi:10.1016/j.arth.2016.02.061
31. Oliver RL. Effect of Expectation and Disconfirmation on Postexposure Product Evaluations: An Alternative Interpretation. *J Appl Psychol*. 1977;62(4):480-486.
32. Haanstra TM. Patients' expectations. Determinants, mechanisms and impact on clinical outcomes. 2015.
33. Tolk JJ, van der Steen M, Janssen RPA, Reijman M. Total knee arthroplasty: what to expect ? A survey of the members of the Dutch Knee Society on long-term recovery after total knee arthroplasty. *J Knee Surg*. 2017;30(6):612-616. doi:10.1055/s-0036-1593868
34. van den Akker-Scheek I, van Raay JJ a M, Reininga IHF, Bulstra SK, Zijlstra W, Stevens M. Reliability and concurrent validity of the Dutch hip and knee replacement expectations surveys. *BMC Musculoskelet Disord*. 2010;11(1):242. doi:10.1186/1471-2474-11-242
35. Lingard EABMM, Katz JNMM, Wright EAP, Sledge CBM, The Kinemax Outcomes Group. Predicting the outcome of total knee arthroplasty. *J Bone Jt Surg - Am Vol*. 2004;86-A(10):2179-2186. doi:86/10/2179 [pii]

36. Singh JA, O'Byrne M, Harmsen S, Lewallen D. Predictors of moderate-severe functional limitation after primary Total Knee Arthroplasty (TKA): 4701 TKAs at 2-years and 2935 TKAs at 5-years. *Osteoarthr Cartil*. 2010;18(4):515-521. doi:10.1016/j.joca.2009.12.001
37. Edwards PK, Mears SC, Lowry Barnes C. Preoperative Education for Hip and Knee Replacement: Never Stop Learning. *Curr Rev Musculoskelet Med*. 2017;10(3):356-364. doi:10.1007/s12178-017-9417-4
38. Husain A, Lee G. Establishing Realistic Patient Expectations Following Total Knee Arthroplasty. *J Am Acad Orthop Surg*. 2015;23(12):707-713. doi:10.5435/JAAOS-D-14-00049
39. Dahm DL, Barnes SA, Harrington JR, Sayeed SA, Berry DJ. Patient-reported activity level after total knee arthroplasty. *J Arthroplasty*. 2008;23(0883-5403; 3):401-407. doi:10.1016/j.arth.2007.05.051
40. Abbasi-Bafghi H, Fallah-Yakhdani HR, Meijer OG, et al. The effects of knee arthroplasty on walking speed: A meta-analysis. *BMC Musculoskelet Disord*. 2012;13(1):66. doi:10.1186/1471-2474-13-66
41. Weiss JM, Noble PC, Conditt MA, et al. What functional activities are important to patients with knee replacements? *Clin Orthop Relat Res*. 2002;(404):172-188. doi:10.1097/01.blo.0000036536.46246.d9
42. Hepinstall MS, Ranawat AS, Ranawat CS. High-flexion total knee replacement: functional outcome at one year. *HSS J*. 2010;6(2):138-144. doi:10.1007/s11420-009-9150-7
43. Jassim SS, Douglas SL, Haddad FS. Athletic activity after lower limb arthroplasty: a systematic review of current evidence. *Bone Joint J*. 2014;96-B(7):923-927. doi:10.1302/0301-620X.96B7.31585
44. Witjes S, Gouttebarge V, Kuijer PPFM, van Geenen RCI, Poolman RW, Kerkhoffs GMMJ. Return to Sports and Physical Activity After Total and Unicondylar Knee Arthroplasty: A Systematic Review and Meta-Analysis. *Sport Med*. 2016;46(2):269-292. doi:10.1007/s40279-015-0421-9
45. Kahlenberg CA, Nwachukwu BU, McLawhorn AS, Cross MB, Cornell CN, Padgett DE. Patient Satisfaction After Total Knee Replacement: A Systematic Review. *HSS J*. 2018:192-201. doi:10.1007/s11420-018-9614-8
46. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16(4):494-502. doi:10.1136/ard.16.4.494
47. Zywiel MG, Mahomed A, Gandhi R, Perruccio A V., Mahomed NN. Measuring expectations in orthopaedic surgery: A systematic review. *Clin Orthop Relat Res*. 2013;471(11):3446-3456. doi:10.1007/s11999-013-3013-8
48. de Groot IB, Favejee MM, Reijman M, Verhaar J a N, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. *Health Qual Life Outcomes*. 2008;6:16. doi:10.1186/1477-7525-6-16
49. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Outcome. *Arthritis Care Res (Hoboken)*. 2011;63 Suppl 1(November):S208-28. doi:10.1002/acr.20632
50. Haverkamp D, Breugem SJM, 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 Orthop*. 2005;76(3):347-352. doi:10.1080/00016470510030814
51. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343.
52. Conner-Spady BL, Marshall DA, Bohm E, et al. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee

replacement. *Qual Life Res*. 2015;24(7):1775-1784. doi:10.1007/s11136-014-0910-6
53. Ruyssen-Witrand A, Fernandez-Lopez CJ, Gossec L, Anract P, Courpied JP, Dougados M. Psychometric properties of the OARSI/OMERACT osteoarthritis pain and functional impairment scales: ICOAP, KOOS-PS and HOOS-PS. *Clin Exp Rheumatol*. 2011;29(2):231-237.
54. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res*. 2002;52(2):69-77. doi:S0022399901002963 [pii]
55. Spinhoven P, Ormel J. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med*. 1997;27(2):363-370.
56. Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess*. 1995;7(4):524-532. doi:10.1037/1040-3590.7.4.524
57. ten Klooster PM., Weekers AM., Eggelmeijer F., et al. Optimisme en / of pessimisme: factorstructuur van de Nederlandse Life Orientation Test Revised. *Psychol Gezondh*. 2010;38(2):89-100.
58. Glaesmer H, Rief W, Martin A, et al. Psychometric properties and population-based norms of the Life Orientation Test Revised (LOT-R). *Br J Health Psychol*. 2012;17(2):432-445. doi:10.1111/j.2044-8287.2011.02046.x
59. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. *J Clin Epidemiol*. 2005;58:595-602. doi:10.1016/j.jclinepi.2004.10.018
60. Bjorgul K, Novicoff WM, Saleh KJ. Evaluating comorbidities in total hip and knee arthroplasty: available instruments. *J Orthop Traumatol*. 2010;11:203-209. doi:10.1007/s10195-010-0115-x
61. Iorio R, Della Valle CJ, Healy WL, et al. Stratification of standardized TKA complications and adverse events: A brief communication knee. *Clin Orthop Relat Res*. 2014;472(1):194-205. doi:10.1007/s11999-013-2980-0
62. Vissers MM, De Groot IB, Reijman M, Bussmann JB, Stam HJ, Verhaar JAN. Functional capacity and actual daily activity do not contribute to patient satisfaction after total knee arthroplasty. *BMC Musculoskelet Disord*. 2010;11(121). doi:10.1186/1471-2474-11-121
63. 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. doi:10.1016/j.arth.2016.11.007
64. Lingard E a, Sledge CB, Learmonth ID. Patient expectations regarding total knee arthroplasty: differences among the United States, United kingdom, and Australia. *J Bone Jt Surg Am*. 2006;88-A(6):1201-1207. doi:10.2106/jbjs.e.00147

Abstract

Introduction

Meeting pre-operative expectations is known to be of major influence on post-operative satisfaction after total knee arthroplasty (TKA). Improved expectation management, resulting in more realistic expectations can potentially lead to higher post-operative satisfaction. Main study objective was o assess the effect of an additional pre-operative education module, with specific attention to realistic expectations for long-term functional recovery, on post-operative satisfaction. We hypothesized that this additional module would lead to a higher postoperative satisfaction rate and a higher postoperative expectation fulfilment

Methods

We conducted a randomized controlled trial with observers blinded to group allocation and patients naïve to study objective. Patients were included between July 2016 and April 2018.Patients were allocated to usual pre-operative education (control group) or usual education plus an additional module on realistic expectations (intervention group). The primary outcome was being very satisfied (numerical rating scale for satisfaction ≥8) with the treatment result at 12-month follow-up. Other outcomes were change in pre-operative expectations measured with the Hospital for Special Surgery Knee Replacement Expectations Survey (HSS-KRES) and postoperative fulfilment of expectations.

Results

A consecutive sample of 459 knee osteoarthritis patients indicated for primary, unilateral TKA were assessed for eligibility. 204 patients were randomized (mean age 68.7 [standard deviation (SD) 9.4] years; 122 women [79.8%]), and 187 patients (91.7%) were available for analysis at follow-up. In the intention-to-treat analysis 69.9% patients were very satisfied with the treatment result in the intervention group, and 58.5% patients in the control group, between group difference 11.4% (95%CI -2.4 – 25.2%). A per-protocol analysis for patients that attended the education session (94.1%) showed 74.4% very satisfied patients in the intervention group and 56.9% in the control group (mean difference 17.4% [95%CI 3.3-31.6%]).

After pre-operative education the expectation scores in the intervention group were significantly lower (mean difference -6.9 [95%CI -10.2 – -3.6]) and unchanged in the control group (mean difference 0.5 [95%CI -2.9 – 3.9). Overall fulfilment of expectations at 12-months follow-up was 70% (SD 28.8) in the intervention group and 58.6% (SD 33.0) in the control group; mean difference 11.4% (95%CI 2.3-20.5).

Conclusion

Improved pre-operative patient education can modify patient expectations, resulting in higher postoperative expectation fulfilment and higher postoperative satisfaction. This is the first randomized controlled trial to confirm the potential of improved expectation management on satisfaction after TKA.

Trial registration

Dutch Trial Registry registration on March 17, 2016. Number: NL5006. https://www.trialregister.nl/trial/5006

chapter 8

The influence of Expectation modification in Knee arthroplasty on Satisfaction of PatiEnts: a randomized Controlled Trial: the EKSPECT Study

J.J. Tolk, R.P.A. Janssen, T.M. Haanstra, M.C. Van der Steen,
S.M.A. Bierma-Zeinstra, M. Reijman

Submitted

Background

Treatment of end-stage knee osteoarthritis with total knee arthroplasty (TKA) generally results in considerable pain reduction and increase in physical function and quality-of-life.[1,2] Nevertheless, approximately one in five primary TKA patients are reported to be not completely satisfied, and only half of patients is reported to be very satisfied with the treatment result.[3–7] The strongest predictor of dissatisfaction after primary TKA is fulfilment of pre-operative expectations on treatment result.[3–6] Therefore, a higher expectation fulfilment rate could potentially lead to higher postoperative satisfaction

Patients have a range of expectations regarding the treatment result of TKA. Most important expectations concern pain relief and physical functioning.[8,9] Pre-operative expectations are often overly optimistic [10,11], and a substantial number of patients has unfulfilled expectations after TKA.[12] Previous research showed that structured pre-operative education on realistic expectations for long term recovery can alter patients' expectations.[13] The effect of expectation management on expectation fulfilment and satisfaction after TKA has not yet been studied.

The primary objective of this randomized controlled trial was to examine whether an additional education module targeted at realistic expectations on relief of symptoms and improvement in physical functioning would increase patient satisfaction 1 year after TKA compared to usual pre-operative education. Secondary analyses were done to determine if the intervention changed pre-operative outcome expectations and postoperative expectation fulfilment. We hypothesized that the additional education module would lead to higher postoperative expectation fulfilment and patient satisfaction scores, when compared to usual pre-operative education.

Methods

We conducted a randomized controlled superiority trial, with a parallel group design and 1:1 allocation ratio. Patients were randomized in 2 groups. Group 1: usual education plus an additional module on realistic expectations for long-term recovery. Group 2: usual education. The study protocol has been reviewed and approved by the Medical Ethics Committee Máxima MC (NL54671.015.15), and was registered in the Dutch Trial Registry (NTR5779). The study design has been published previously and has not been altered or amended.[14]

All patients presenting at the Máxima MC between July 2016 and May 2018 with symptomatic and radiographic knee osteoarthritis indicated for primary TKA were assessed for eligibility by a research nurse. Patients meeting one or more of the following criteria were excluded from participation: presence of a medical illness that results in life expectancy shorter than 1-year, previous contralateral TKA, unicompartmental knee arthroplasty, staged or bilateral TKA, insufficient command of the Dutch language or legally incompetent adults. If a patient was eligible and willing to participate, informed consent was signed.

Group allocation took place by receiving a computer-generated (Research Manager version 5.30.0.6, Cloud9 Software, Deventer, the Netherlands) randomization number from the coordinating investigator. Because information on outcome after TKA provided by each orthopaedic surgeon during the consultation might be different on details, block randomization with variable size of blocks, stratified for the 5 orthopaedic surgeons treating the patients was used to account for potential information bias. All staff was blinded for allocation during study procedures and data analysis. Patients were naïve to study objective and difference in education between the intervention and control group. Only after data analysis, unblinding took place.

Intervention

The intervention under study is a joint-specific educational module aimed at achieving realistic expectations on long-term recovery after TKA.[14] For the intervention group this module was added to the existing pre-operative education program as described for the control group. The additional 30-minutes module on expectations for long-term recovery addresses a set of expectations that are most important to TKA patients.[8,9,12] Development and content of the module are explicated in the study design paper.[14] The evidence-based module describes the average treatment result for a primary TKA patient regarding pain reduction, physical function, performing social activities and psychological well-being one year after surgery. Furthermore, individual factors that influence treatment result, such as age, co-morbidity, body mass index (BMI), psychosocial factors, pain severity and pre-operative functional status are addressed.[15,16] In summary, the education module states that improvement can be expected in domains of pain, function, activities, and psychological well-being but return to normal is not likely to occur. Limitations should predominantly be expected for more demanding physical activities and sports.

Patients in the control arm received the standard 90-minute multidisciplinary pre-operative education program. This program educates patients on what to expect for the perioperative period, but does not include information on long-term recovery. This module currently is the standard care in Máxima MC for all TKA patients.

Outcome measurement

The primary outcome was whether the patient is very satisfied with the treatment result at 12-months follow-up. Satisfaction was measured using a self-reported numeric rating scale (NRS) for satisfaction with treatment result, ranging from 0 (very dissatisfied) to 10 (very satisfied).[17] 'Very satisfied with treatment result' was defined as an NRS satisfaction score of $^{3}8$.[14] All questionnaires were completed digitally and study data was stored coded using data management software (Research Manager version 5.30.0.6, Cloud9 Software, Deventer, the Netherlands).

At baseline the following patient characteristics were obtained: age, sex, side affected and body mass index (BMI).[18]

Patient expectations were measured at baseline (after inclusion but before pre-operative education) and at admission for surgery (after the pre-operative education) using the Dutch version of the Hospital for Special Surgery Knee Replacement Survey (HSS-KRES).[19] The HSS-KRES is a 19-item, reliable and valid self-administered survey measuring probability-based

outcome expectations in TKA patients (range 0 [lowest] to 100 [highest]).[8,19,20] To assess to what extent expectations had been fulfilled 12 months postoperatively, the expectation fulfilment version of the HSS-KRES was used.[21] This score compares the perceived actual postoperative outcome to the pre-operative score for each HSS-KRES item.[21] For items scored as "not applicable" in either the pre- or postoperative questionnaire a fulfilment score could not be calculated. Frequencies of unfulfilled, fulfilled or exceeded expectations were calculated for each item. A total score representing the percentage of items with fulfilled or exceeded expectations compared to the total number of items with a fulfilment score was calculated.

Pain scores, physical function and general health status were measured at baseline and 12-month follow-up using patient reported outcome measures. Physical function was assessed using the Dutch version of the Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form (KOOS-PS [(range 0 [worst]-100 [best).[22,23] General health status was measured using the Dutch version of the EuroQol 5D-3L (EQ-5D) descriptive index (range 1 [best possible health] through 0 [death] to −0.59 [worse than death]), and EQ-5D health scale (range, 0 [worst] to 100 [best]). [24,25] Pain was measured using an NRS score for pain during activity and at rest (NRS pain [range 0 (no pain) to 10 (worst imaginable pain)].[26,27] Anxiety and depressive symptoms were assessed pre-operatively with the Dutch version of the Hospital Anxiety and Depression Scale (HADS [range 0 (no symptoms) to 21 (severe symptoms), participants with a score ³8 were defined as having anxiety or depressive symptoms[28,29]

Sample size calculations

Sample size was based on the hypothesis that more patients in the intervention group would be very satisfied with treatment outcome one year postoperative. Previous work has shown that approximately 50% of patients is very satisfied after TKA.[7] Power was set at 80% and an alpha of 0.05 was used. To increase the proportion of very satisfied patients from 50% of to 70% would require 90 patients in each group. Accounting for 15% potential dropout rate, the sample size required was 204 patients.

Statistical analysis

For descriptive statistics continuous variables were presented as means with standard deviations (SD) or median with interquartile range (IQR) depending on data distribution, and for discrete variables counts and percentages.

The primary analysis was performed according to intention-to-treat principles. A secondary per-protocol-analysis was performed addressing patients that actually received the allocated intervention.

Primary outcome was analysed using logistic regression analyses with percentage of patients being very satisfied with the treatment result as dependent and intervention as independent variable. Variables that are known to be associated with patient satisfaction, were considered as covariates in the primary analysis; age, gender, BMI, baseline expectations, anxiety and depression. The "change-in-estimate" approach was used; covariates that changed the estimate of the causal effect for the group allocation by more than 10% were selected.[30] Predefined secondary outcome analyses were performed

difference between intervention and control group on change in HSS-KRES outcome expectations total scores before and after pre-operative education (baseline vs pre-operative), and fulfilment of pre-operative expectations. Linear regression analyses were performed, using the same covariates as for the primary analysis. In comparing change in expectations on individual HSS-KRES items after the pre-operative education session and fulfilment of the expectations on individual items we accounted for multiple testing. The threshold for significance on these items was set as $p<0.003$ as there are 19 items on the HSS knee replacement expectations scale. Follow-up pain scores, physical function and general health status were compared using a t-test or Mann-Whitney-U test depending on normality of data distribution. Statistical calculations were made using IBM® SPSS software, version 24.0.

Results

Study population

A total of 459 primary TKA patients were screened for eligibility between July 2016 and April 2018. 204 patients were eligible for participation and willing to participate (figure 1). 101 patients were randomly allocated to the intervention group and 103 to the control group. The number of patients was not equal, because of the block randomization method used. Reasons for not participating, attendance of the pre-operative education module and reasons for loss-to-follow-up are summarized in figure 1. Patient characteristics of both groups are shown in Table 1, no clinically meaningful imbalances were observed.

Primary outcome

Satisfaction with treatment result

In the intention-to-treat analysis the frequency of the primary outcome (very satisfied with the treatment result) was 65/93 (69.9%) patients in the intervention group, and 55/94 (58.5%) patients in the control group. This difference was not statistically significant; between group difference 11.4% (95%CI -2.4 – 25.2%). None of the co-variates that were considered (age, gender, BMI, baseline expectations, anxiety and depression) resulted in > 10% change in the estimate. Therefore, it was concluded that the co-variates did not introduce bias in the interpretation of the primary outcome and an uncorrected analysis was performed.

Not all patients attended the pre-operative education sessions. Of the patients that were available for analysis at follow-up in the intervention group 86 patients had received the allocated intervention, in the control group 86 patients had received the allocated intervention (figure 1). There was no cross-over between groups. For the patients that did attend the meetings a per-protocol analysis was performed. In this analysis the number of patients that was very satisfied with the treatment result was significantly higher in the intervention group in comparison with the control group; 61/86 (74.4%) versus 49/86 (56.9%) respectively (mean difference 17.4% [95%CI 3.3-31.6%]).

Secondary outcome

Patient reported outcome measures

At 12-month follow-up pain scores, patient reported physical function and health status did not show significant differences between the two groups (Table 2).

Expectations and expectation fulfilment

Baseline expectation total scores were similar between groups; mean HSS-KRES score 69.1 (SD 18.2) for the intervention group and 72.4 (17.6) for the control group (mean difference -3.3 [95%CI -8.2 – 1.7]), as were individual item scores (supplementary file 1). After the intervention the HSS-KRES expectation total scores in the control group remained unchanged (mean difference 0.5 [95%CI -2.9 – 3.9], p=0.777) whereas expectations total scores were significantly lower in the intervention group (mean difference -6.9 [95%CI -10.2 – -3.6], p<0.001). The baseline, pre-operative and follow-up HSS-KRES scores for all individual items are shown in supplementary file 1.

In the control group none of the HSS-KRES items showed significant change in scores after the pre-operative education, compared to baseline scores (supplementary file 2). In the intervention group scores significantly changed on 5 of the 19 HSS-KRES items compared to the baseline scores in the intervention group, accounting for multiple testing. Expectations regarding improvement in ability to walk medium and long distances, kneeling, squatting and driving were lower compared to baseline in the intervention group.

Overall fulfilment of expectations was significantly different between groups. In the intervention group the overall expectation fulfilment rate was 70% (SD 28.8) compared to 58.6% (SD (33.0) in the control group; mean difference 11.4% (95%CI 2.3-20.5). The fulfilment rate for the individual items of the HSS-KRES scores is presented in supplementary file 2. At 12-month follow-up only expectations for improvement in the ability to kneel was significantly more often fulfilled or exceeded in the intervention group.

Table 1. Patient characteristics

	Intervention group (n=101)	**Control group (n=103)**
Age, years	68.4 (8.7)	69.0 (10.1)
Male sex, n (%)	41 (40.6)	41 (39.8)
BMI, kg/m²	29.2 (5.2)	28.7 (4.8)
NRS pain, median (IQR) - *at rest* - *during activity*	 5 (3-7) 8 (8-9)	 5 (3-7) 8 (7-9)
EQ-5D, median (IQR) - *health scale* - *descriptive index*	 73 (50-82.5) 0.65 (0.22-0.78)	 78 (63-85) 0.69 (0.42-0.78)
KOOS-PS, median (IQR)	54.4 (46.1-62.0)	51.2 (46.1-62.0)
HADS, n (%) - *depression score ³ 8* - *anxiety score ³ 8*	 16 (15.8) 15 (14.9)	 13 (12.6) 12 (11.7)
Time from baseline to intervention, days, median (IQR)	52 (37-75)	49 (34-63)
Time from intervention to surgery, days, median (IQR)	16 (8-24)	18 (7-26)

Continuous data are presented as mean and standard deviation between parentheses or reported otherwise as mentioned depending on normality of data distribution.
IQR; interquartile range, BMI; Body Mass Index, NRS; Numerical Rating Scale, KOOS-PS; Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form, HADS; Hospital Anxiety and Depression Scale.

Discussion

This RCT showed that in patients indicated for TKA additional education on realistic outcome expectations resulted in a smaller difference in very satisfied patients between groups than expected. Not in the intention-to-treat analysis, but only in the per-protocol analyses was the percentage very satisfied patients significantly higher in the intervention group. Furthermore, the intervention resulted in modification of pre-operative expectations and a higher postoperative expectation fulfilment rate, compared to the control group.

The mean satisfaction scores as well as the number of patients that were very satisfied with the treatment result were higher in the intervention group, compared to the control group. However, this difference was statistically significant only in the per-protocol analysis. Partly this can be explained by the relatively high proportion of patients being very satisfied in the control group, which was almost 60% whereas we based our power calculations on 50% of patients being very satisfied.[7] This diminished the contrast and thus could have resulted in the study being slightly underpowered. On the other hand, in the predefined per-protocol analysis the difference was statistically significant. In this analysis we evaluated only patients that did visit the pre-operative education class (in either the intervention or control group). In our view, this result supports the effect of the intervention; the education class obviously only had an impact on expectations of patients that actually attended the meeting. Therefore, we conclude that the intervention resulted in a higher proportion of patients being very satisfied with the treatment result at 1-year follow-up.

In the intervention group significantly lower expectations were observed after the education module. This result is in line with a RCT by Mancuso et al., who evaluated the effect of a comparable intervention.[13] In the analysis of the specific HSS-KRES expectation items , the items that showed a significant reduction in expectation score largely coincide with expectations that previously have been reported to be the items with the most unfulfilled expectations posteopratively.[12] Therefore it seems likely that the intervention not only modified pre-operative expectations, but was actually able to specifically modify pre-operative expectation that were unrealistic. This conclusion is supported by the finding that in the intervention group the overall expectation fulfilment rate was considerably higher.

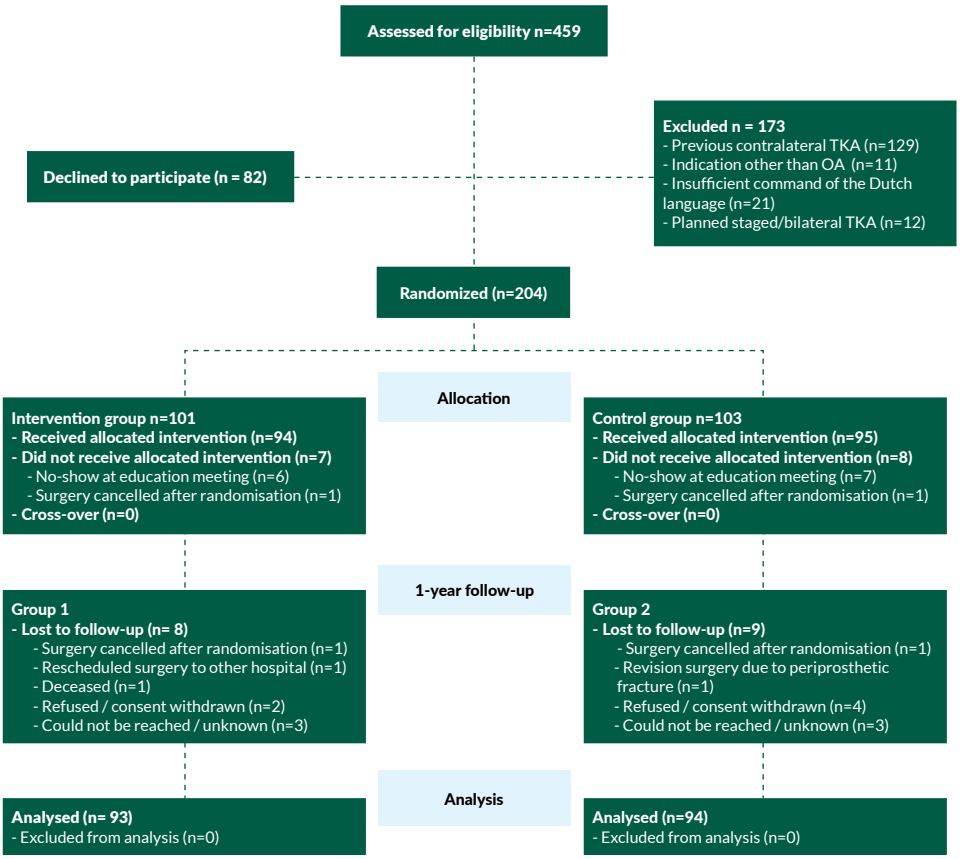
Table 2. 12-month follow-up patient reported outcome scores

	Intervention group (n=93)	Control group (n=94)	p-value
NRS satisfaction score	8.0 (7-9)	8.0 (6.8-9)	0.056
NRS pain - at rest - during activity	 1.0 (0-2) 2.0 (1-4)	 1.0 (0-2.75) 2.0 (1.0-4.75)	 0.728 0.958
EQ-5D - health scale - descriptive index	 79 (70-86) 0.84 (0.77-1.00)	 80 (70-85) 0.81 (0.78-1.00)	 0.948 0.678
KOOS-PS	28.6 (21.2-35.3)	27.5 (18.6-37.0)	0.727

Data are presented as median and interquartile range between brackets. Statistical significance using Mann-Whitney-U test unless otherwise mentioned. NRS; numerical rating scale, KOOS-PS; Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form.

This supports the hypothesis that the expectation modification intervention resulted in more realistic expectations in the TKA patients under study.

There has been some debate on the goal of pre-operative expectation management.[31,32] TKA patients with higher expectations are reported to have better outcome in terms of pain reduction and self-reported functional outcome, but the level of pre-operative expectations is not related to satisfaction.[31] Postoperative satisfaction is highly correlated to fulfilment of pre-operative expectations, independent from the treatment actual outcome in terms of pain, quality-of-life and function.[4,33] The data from the present study shows the same correlation between expectation fulfilment and satisfaction (data not shown). On the other



hand, patients' expectations for TKA patients are recommended to be as high as possible.[31] This would allow patients to fully benefit from positive effects of anxiety reduction, better treatment adherence and beneficial coping mechanisms that are associated with higher expectations.[34,35] Furthermore, too low expectations might result in a nocebo effect and less motivation to obtain full benefit from the surgery.[36,37] In the present study the mean expectations scores were lowered by the intervention, but this was not a goal per se. Probably some patients should have their expectations increased to obtain the advantage in treatment outcome.[31] To identify and specifically address these subgroups further research is warranted.

The present study had several strengths. One of the major strengths is that, despite the great amount of attention given to expectation management in recent orthopaedic literature, this is the first study to translate the theoretical advantage of improved expectation management to a clinically useful intervention. The adequate power and relatively small proportion of patients lost to follow-up substantiates the conclusion of this study. Furthermore, in previous research on expectation management poor measurement methods and inconsistent definition of constructs under study are common.[38,39] It this study the construct measured and the theoretical framework it derives from was adequately defined,[14,39] adding to the interpretability of the results. Finally, the nature of the intervention presented allows for easy integration in current clinical processes. Of course, this is dependent on the pre-operative education logistics in each hospital, but the limited burden for patients and staff allows for broader application beyond this study setting.

Some limitations of the study have to be addressed. First, a limitation of the study was that the group-based approach of the intervention only allowed for limited personalization of the expectation management. It would be useful to be able to identify patients at risk of having unrealistic expectations, and accurately predict the most likely treatment result to further individualize expectation management. Efforts in this regard have been made,[40,41] but clinical implementation still has to be evaluated. Another limitation is that only osteoarthritis patients indicated for primary TKA in the Dutch clinical setting were included. Expectations and satisfaction rates are known to differ across populations and for indications for TKA other than primary knee OA.[6,42] Therefore the results might not be fully representative for other countries and revision or unicondylar knee arthroplasty patients. These factors might limit the generalizability of the study results to some extent. Although it is not likely that this affects the concept of achieving realistic expectations to improve patient satisfaction.

Conclusion

Improved pre-operative patient education can modify patient expectations, resulting in higher postoperative expectation fulfilment and higher postoperative satisfaction. This indicates we should motivate patients to visit these pre-operative information meetings. This is the first prospective, randomized controlled trial to confirm the potential of improved expectation management on patient satisfaction in TKA patients.

Supplementary file 1. Hospital for Special Surgery Knee Replacement Expectations Survey scores									
	Baseline HSS-KRES			Preop HSS-KRES			HSS-KRES fulfilment score		
	Intervention group	Control group	p-value*	Intervention group	Control group	p-value*	Intervention group	Control group	p-value*
Q1 Relieve pain	3.3 (0.6)	3.4 (0.6)	0.048	3.1 (0.7)	3.4 (0.6)	<0.001	2.9 (0.9)	3.1 (1.0)	0.137
Q2 Walk short distance	3.4(0.6)	3.5 (0.6)	0.115	3.3 (0.8)	3.4 (0.6)	0.231	3.0 (1.2)	3.0 (1.0)	0.592
Q3 Walk medium distance	3.2 (0.9)	3.3 (0.7)	0.134	2.8 (0.9)	3.2 (0.7)	0.001	2.7 (1.2)	2.9 (1.1)	0.225
Q4 Walk long distance	2.7 (1.1)	3.0 (0.8)	0.050	2.4 (1.1)	3.0 (0.9)	<0.001	2.3 (1.4)	2.5 (1.2)	0.207
Q5 Remove need for cane	2.5 (1.8)	2.9 (1.7)	0.127	2.6 (1.7)	3.0 (1.6)	0.074	2.8 (1.6)	2.7 (1.7)	0.491
Q6 Make knee straight	3.2 (0.9)	3.3 (1.0)	0.287	3.1 (1.10	3.3 (1.0)	0.160	3.2 (1.1)	3.3 (1.1)	0.659
Q7 Go up stairs	3.1 (1.0)	3.3 (0.8)	0.169	2.9 (1.1)	3.3 (0.8)	0.003	2.6 (1.2)	2.7 (1.2)	0.571
Q8 Go down stairs	3.1 (0.9)	3.2 (0.9)	0.241	2.9 (1.1)	3.3 (0.9)	0.007	2.6 (1.2)	2.6 (1.3)	0.887
Q9 Kneel	2.6 (1.1)	2.6 (1.0)	0.905	1.9 (1.1)	2.7 (1.0)	<0.001	1.6 (1.2)	1.7 (1.3)	0.284
Q10 Squat	2.5 (1.1)	2.6 (1.1)	0.509	1.9 (1.0)	2.5 (1.1)	<0.001	1.8 (1.2)	1.8 (1.3)	0.810
Q11 Use public transportation or drive	2.9 (1.3)	3.0 (1.5)	0.450	2.4 (1.6)	3.0 (1.4)	0.005	2.8 (1.5)	2.7 (1.6)	0.515
Q12 Be employed	1.2 (1.7)	1.4 (1.8)	0.434	1.1 (1.6)	1.1 (1.7)	0.993	1.6 (1.8)	1.4 (1.8)	0.604
Q13 Recreational activities	2.7 (1.3)	3.0 (1.2)	0.178	2.6 (1.3)	3.0 (1.2)	0.025	2.4 (1.4)	2.4 (1.5)	0.985
Q14 Daily activities	3.1 (0.9)	3.3 (0.9)	0.104	3.0 (1.2)	3.3 (1.0)	0.011	2.7 (1.3)	3.0 (1.1)	0.159
Q15 Sports	2.4 (1.4)	2.7 (1.1)	0.101	2.1 (1.3)	2.8 (1.2)	<0.001	2.1 (1.4)	2.3 (1.3)	0.161
Q16 Sitting to standing	3.2 (0.9)	3.3 (0.9)	0.212	2.9 (1.2)	3.5 (0.7)	<0.001	2.7 (1.2)	2.9 (1.1)	0.245
Q17 Interact with others	2.9 (1.2)	2.9 (1.1)	0.537	2.6 (1.3)	3.1 (1.1)	0.009	2.6 (1.3)	2.5 (1.4)	0.608
Q18 Sexual activity	2.2 (1.7)	1.8 (1.8)	0.163	2.1 (1.7)	2.1 (1.7)	0.992	1.7 (1.6)	1.7 (1.7)	0.883
Q19 Psychological well-being	2.6 (1.6)	2.5 (1.8)	0.454	2.2 (1.6)	2.5 (1.7)	0.265	2.5 (1.6)	2.4 (1.6)	0.699
HSS-KRES total score	69.1 (18.2)	72.4 (17.6)	0.19	62.8 (19.2)	72.8 (16.2)	<0.001	62.9 (23.0)	61.9 (25.9)	0.80

*Hospital for Special Surgery Knee Replacement Expectations Survey (HSS-KRES) for each question at baseline and pre-operative (after the education but before surgery). Pre-operative HSS-KRES expectations fulfilment score at 12-month follow-up. Data presented as mean with standard deviation between brackets. *T-test for independent samples. Q; question.*

Supplementary file 2. Change in expectations after pre-operative education and postoperative expectation fulfilment						
	Change in HSS-KRES score after pre-operative education, mean (SD)			Expectations fulfilled /exceeded, n/N (%)		
	Intervention group (n=101)	Control group (n=103)	p-value†	Intervention group (n=93)	Control group (n=94)	p-value‡
Q1 Relieve pain	-0.2 (0.7)	0.0 (0.7)	0.067	67/88 (76.1)	61/91 (67)	0.177
Q2 Walk short distance	-0.1 (0.9)	-0.1 (0.8)	0.812	61/86 (70.9)	57/90 (63.3)	0.284
Q3 Walk medium distance	-0.4 (0.9) *	-0.1 (0.9)	0.064	60/84 (71.4)	55/88 (62.5)	0.214
Q4 Walk long distance	-0.4 (1.2) *	-0.1(0.9)	0.027	56/71 (78.9)	49/84 (58.3)	0.006
Q5 Remove need for cane	0.1 (2.2)	0.1 (1.7)	0.787	45/53 (84.9)	43/59 (72.9)	0.122
Q6 Make knee straight	-0.1 (1.0)	0.0 (1.1)	0.390	62/83 (74.7)	65/84 (77.4)	0.685
Q7 Go up stairs	-0.2 (1.0)	0.1 (0.8)	0.065	51/82 (62.2)	45/85 (52.9)	0.227
Q8 Go down stairs	-0.2 (1.0)	0.0 (0.8)	0.078	52/81 (64.2)	44/83 (53.0)	0.146
Q9 Kneel	-0.7 (1.1) *	0.1 (1.1)	<0.001	44/60 (73.3)	28/73 (38.4)	<0.001
Q10 Squat	-0.6 (1.1) *	-0.1 (1.1)	<0.001	44/67 (65.7)	33/73 (45.2)	0.015
Q11 Use public transportation or drive	-0.5 (1.5) *	0.0 (1.7)	0.041	44/59 (74.6)	47/68 (69.1)	0.496
Q12 Be employed	-0.2 (1.3)	-0.3 (1.7)	0.511	17/22 (77.3)	14/21 (66.7)	0.438
Q13 Recreational activities	-0.2 (1.4)	0.0 (1.3)	0.266	46/67 (68.7)	38/72 (52.8)	0.560
Q14 Daily activities	-0.2 (1.2)	0.0 (1.1)	0.139	49/77 (63.6)	52/86 (60.5)	0.677
Q15 Sports	-0.3 (1.4)	0.1 (1.1)	0.014	49/63 (77.8)	46/77 (59.7)	0.023
Q16 Sitting to standing	-0.3 (1.2)	0.2 (0.9)	0.003	42/78 (53.8)	49/88 (55.7)	0.813
Q17 Interact with others	-0.3 (1.4)	0.1 (1.3)	0.031	51/72(70.8)	41/75 (45.7)	0.043
Q18 Sexual activity	-0.1 (1.5)	0.2 (1.6)	0.125	30/42 (71.4)	27/43 (62.8)	0.397
Q19 Psychological well-being	-0.4 (1.7)	0.0 (1.9)	0.077	40 /52 (76.9)	34/58 (58.6)	0.041

*Change in mean Hospital for Special Surgery Knee Replacement Expectations Survey (HSS-KRES) score between baseline and after pre-operative education, and number of patients with pre-operative expectations that were fulfilled or exceeded at 1-year follow-up. The number of the patients (n) that had the respective KSS-KRES item fulfilled or exceeded, within the group of patients (N) that had pre- and postoperative scores for that HSS-KRES item. SD; Standard Deviation, Q; question. *statistically significant different from baseline expectation score † Independent Samples T-test ‡ Pearsons chi-square test. Significance set at p<0.003 to account for multiple testing for each of the 19 items.*

CONSORT CHECKLIST

Table. CONSORT 2010 Checklist of Information to Include When Reporting a Randomized Trial[a]

Section and Topic	Item No.	Checklist Item	Reported on Page No.
Title and abstract			
	1a	Identification as a randomized trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	3-6
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	7
	2b	Specific objectives or hypotheses	7
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	7
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	8
Participants	4a	Eligibility criteria for participants	8
	4b	Settings and locations where the data were collected	8
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	9
Outcomes	6a	Completely defined prespecified primary and secondary outcome measures, including how and when they were assessed	10-11
	6b	Any changes to trial outcomes after the trial commenced, with reasons	8
Sample size	7a	How sample size was determined	11
	7b	When applicable, explanation of any interim analyses and stopping guidelines	NA
Randomization			
Sequence generation	8a	Method used to generate the random allocation sequence	8,9
	8b	Type of randomization; details of any restriction (such as blocking and block size)	8,9
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	8,9
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	8,9
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	8,9
	11b	If relevant, description of the similarity of interventions	9
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	12,13
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	12
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome	Figure 1
	13b	For each group, losses and exclusions after randomization, together with reasons	Figure 1
Recruitment	14a	Dates defining the periods of recruitment and follow-up	13
	14b	Why the trial ended or was stopped	13
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1
Numbers analyzed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	13,14
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	13-15
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	13-15
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing prespecified from exploratory	13-15
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	NA
Comment			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	18
Generalizability	21	Generalizability (external validity, applicability) of the trial findings	18
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	15-18
Other information			
Registration	23	Registration number and name of trial registry	6,8
Protocol	24	Where the full trial protocol can be accessed, if available	8
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	NA

[a] We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomized trials, noninferiority and equivalence trials, nonpharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up-to-date references relevant to this checklist, see http://www.consort-statement.org.

References

1. Ruiz D, Koenig L, Dall TM, et al. The Direct and Indirect Costs to Society of Treatment for End-Stage Knee Osteoarthritis. *J Bone Jt Surg Am*. 2013;95(16):1473-1480. doi:10.2106/JBJS.L.01488
2. Tolk JJ, Janssen RPA, Prinsen CAC, et al. The OARSI core set of performance-based measures for knee osteoarthritis is reliable but not valid and responsive. *Knee Surgery, Sport Traumatol Arthrosc*. 2019;27(9):2898-2909. doi:10.1007/s00167-017-4789-y
3. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: Who is satisfied and who is not? *Clin Orthop Relat Res*. 2010;468(1):57-63. doi:10.1007/s11999-009-1119-9
4. Hamilton DF, Lane J V, Gaston P, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open*. 2013;3(4):1-8. doi:10.1136/bmjopen-2012-002525
5. Noble PC, Conditt M a, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452(452):35-43. doi:10.1097/01.blo.0000238825.63648.1e
6. 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):148-152. doi:10.1302/0301-620x.95b11.32767
7. Vissers MM, De Groot IB, Reijman M, Bussmann JB, Stam HJ, Verhaar JAN. Functional capacity and actual daily activity do not contribute to patient satisfaction after total knee arthroplasty. *BMC Musculoskelet Disord*. 2010;11(121). doi:10.1186/1471-2474-11-121
8. Mancuso CA, Sculco TP, Wickiewicz TL, Jones EC, Robbins L, Warren RF. Patients' expectations of knee surgery. *J Bone Jt Surg Am*. 2001;83-A(7):1005-1012.
9. Smith EJ, Soon V-L, Boyd A, McAllister J, Deakin A, Sarungi M. What do Scottish Patients Expect of Their Total Knee Arthroplasty? *J Arthroplasty*. 2016;31(4):786-792. doi:10.1016/j.arth.2015.10.021
10. Mannion AF, Kämpfen S, Munzinger U, Kramers-de Quervain I. The role of patient expectations in predicting outcome after total knee arthroplasty. *Arthritis Res Ther*. 2009;11(5):R139. doi:10.1186/ar2811
11. 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 Orthop*. 2009;80(1):55-61. doi:10.1080/17453670902805007
12. Tilbury C, Haanstra TM, Leichtenberg CS, et al. Unfulfilled Expectations After Total Hip and Knee Arthroplasty Surgery: There Is a Need for Better Preoperative Patient Information and Education. *J Arthroplasty*. 2016;31(10):2139-2145. doi:10.1016/j.arth.2016.02.061
13. Mancuso C, Graziano S, Briskie L, et al. Randomized trials to modify patients' preoperative expectations of hip and knee arthroplasties. *Clin Orthop Relat Res*. 2008;466(2):424-431. doi:10.1007/s11999-007-0052-z
14. Tolk JJ, Janssen RPA, Haanstra TM, Bierma-zeinstra SMA, Reijman M. The EKSPECT study : the influence of Expectation modification in Knee arthroplasty on Satisfaction of PatiEnts : study protocol for a randomized Controlled Trial. *Trials*. 2018;19(437).
15. Lingard EABMM, Katz JNMM, Wright EAP, Sledge CBM, The Kinemax Outcomes Group. Predicting the outcome of total knee arthroplasty. *J Bone Jt Surg - Am Vol*. 2004;86-A(10):2179-2186. doi:86/10/2179 [pii]
16. Singh JA, O'Byrne M, Harmsen S, Lewallen D. Predictors of moderate-severe functional limitation after primary Total Knee Arthroplasty (TKA): 4701 TKAs at 2-years and 2935 TKAs at 5-years. *Osteoarthr Cartil*. 2010;18(4):515-521. doi:10.1016/j.joca.2009.12.001

17. Kahlenberg CA, Nwachukwu BU, McLawhorn AS, Cross MB, Cornell CN, Padgett DE. Patient Satisfaction After Total Knee Replacement: A Systematic Review. *HSS J*. 2018:192-201. doi:10.1007/s11420-018-9614-8
18. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. 1957;16(4):494-502. doi:10.1136/ard.16.4.494
19. van den Akker-Scheek I, van Raay JJ a M, Reininga IHF, Bulstra SK, Zijlstra W, Stevens M. Reliability and concurrent validity of the Dutch hip and knee replacement expectations surveys. *BMC Musculoskelet Disord*. 2010;11(1):242. doi:10.1186/1471-2474-11-242
20. Zywiel MG, Mahomed A, Gandhi R, Perruccio A V., Mahomed NN. Measuring expectations in orthopaedic surgery: A systematic review. *Clin Orthop Relat Res*. 2013;471(11):3446-3456. doi:10.1007/s11999-013-3013-8
21. Tilbury C, Haanstra TM, Leichtenberg CS, et al. Unfulfilled Expectations After Total Hip and Knee Arthroplasty Surgery: There Is a Need for Better Preoperative Patient Information and Education. *J Arthroplasty*. 2016;31(10):2139-2145. doi:10.1016/j.arth.2016.02.061
22. de Groot IB, Favejee MM, Reijman M, Verhaar J a N, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. *Health Qual Life Outcomes*. 2008;6:16. doi:10.1186/1477-7525-6-16
23. Singh JA, Luo R, Landon GC, Suarez-Lmazor ME. Reliability and Clinically Important Improvement Thresholds for Osteoarthritis Pain and Function scales: A Multicenter study. *J Rheumatol*. 2014;41(2):509-515. doi:10.1038/jid.2014.371
24. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343.
25. Paulsen A, Roos EM, Pedersen AB, Overgaard S. Minimal clinically important improvement (MCII) and patient-acceptable symptom state (PASS) in total hip arthroplasty (THA) patients 1 year postoperatively. *Acta Orthop*. 2014;85(1):39-48. doi:10.3109/17453674.2013.867782
26. Ruyssen-Witrand A, Fernandez-Lopez CJ, Gossec L, Anract P, Courpied JP, Dougados M. Psychometric properties of the OARSI/OMERACT osteoarthritis pain and functional impairment scales: ICOAP, KOOS-PS and HOOS-PS. *Clin Exp Rheumatol*. 2011;29(2):231-237.
27. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain*. 2004;8(4):283-291. doi:10.1016/j.ejpain.2003.09.004
28. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res*. 2002;52(2):69-77. doi:S0022399901002963 [pii]
29. Spinhoven P, Ormel J. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med*. 1997;27(2):363-370.
30. Greenland S. Invited commentary: variable selection versus shrinkage in the control of multiple confounders. *Am J Epidemiol*. 2008;167:523-529.
31. Jain D, Nguyen L-CL, Bendich I, et al. Higher patient expectations predict higher patient-reported outcomes, but not satisfaction, in total knee arthroplasty patients: a prospective multicenter study. *J Arthroplasty*. 2017;32(9S):S166-S170. doi:10.1016/j.arth.2017.01.008
32. 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. doi:10.1016/j.arth.2017.07.021
33. Culliton SE, Bryant DM, MacDonald SJ, Hibbert KM, Chesworth BM. Effect of an e-Learning Tool on Expectations and Satisfaction Following Total Knee Arthroplasty: A Randomized Controlled Trial. *J Arthroplasty*. 2018;33(7):2153-2158. doi:10.1016/j.arth.2018.02.040

34. Flood AB, Lorence DP, Ding J, McPherson K, Black NA. The role of expectations in patients' reports of post-operative outcomes and improvement following therapy. *Med Care*. 1993;31(11):1043-1056. doi:10.1097/00005650-199311000-00006
35. Colloca L, Miller FG. Harnessing the placebo effect: the need for translational research. *Philos Trans R Soc Lond B Biol Sci*. 2011;366(1572):1922-1930. doi:10.1098/rstb.2010.0399
36. Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H. The role of expectancies in the placebo effect and their use in the delivery of health care: A systematic review. *Health Technol Assess (Rockv)*. 1999;3(3):1-96. doi:10.3310/hta3030
37. Ebrahim S, Malachowski C, Kamal el Din M, et al. Measures of Patients' Expectations About Recovery: A Systematic Review. *J Occup Rehabil*. 2014;25(1):240-255. doi:10.1007/s10926-014-9535-4
38. Haanstra TM, van den Berg T, Ostelo RW, et al. Systematic review: do patient expectations influence treatment outcomes in total knee and total hip arthroplasty? *Health Qual Life Outcomes*. 2012;10(1):152. doi:10.1186/1477-7525-10-152
39. Tolk JJ, Haanstra TM, Reijman M. Letter to the editor on "What so Scottish patients expect of their total knee arthroplasty?" *J Arthroplasty*. 2016;31(10):2374. doi:10.1016/j.arth.2016.05.023
40. Tolk JJ, Janssen RPA, Haanstra TM, van der Steen MMC, Bierma Zeinstra SMA, Reijman M. Outcome Expectations of Total Knee Arthroplasty Patients: The Influence of Demographic Factors, Pain, Personality Traits, Physical and Psychological Status. *J Knee Surg*. 2019;1(212). doi:10.1055/s-0039-1692632
41. Tolk JJ, Waarsing J (Erwin) H, Janssen RPA, van Steenbergen LN, Bierma-Zeinstra SMA, Reijman M. Development of Preoperative Prediction Models for Pain and Functional Outcome After Total Knee Arthroplasty Using The Dutch Arthroplasty Register Data. *J Arthroplasty*. 2019:Epub ahead of print. doi:10.1016/j.arth.2019.10.010
42. Lingard E a, Sledge CB, Learmonth ID. Patient expectations regarding total knee arthroplasty: differences among the United States, United kingdom, and Australia. *J Bone Jt Surg Am*. 2006;88-A(6):1201-1207. doi:10.2106/jbjs.e.00147

Abstract

Background

One of the main determinants of treatment satisfaction after total knee arthroplasty (TKA) is the fulfilment of preoperative expectations. For optimal expectation management, it is useful to accurately predict the treatment result. Multiple patient factors registered in the Dutch Arthroplasty Register (LROI) can potentially be utilized to estimate the most likely treatment result. The aim of the present study was to create and validate models that predict residual symptoms for patients undergoing primary TKA for knee osteoarthritis.

Methods

Data was extracted from the LROI of all TKA patients who had pre- and postoperative PROMs registered. Multivariable logistic regression analyses were performed to construct predictive algorithms for satisfaction, treatment success, and residual symptoms concerning pain at rest and during activity, sit-to-stand movement, stair negotiation, walking, performance of activities of daily living, kneeling and squatting. We assessed predictive performance by examining measures of calibration and discrimination.

Results

Data of 7071 patients could be included for data analysis. Residual complaints on kneeling (female 72% / male 59%) and squatting (female 71% /male 56%) were reported most frequently, and least residual complaints were scored for walking (female 16% / male 12%) and pain at rest (female 18% / male 14%). The predictive algorithms were presented as clinical calculators that present the probability of residual symptoms for an individual patient. The models for residual symptoms concerning sit-to-stand movement, stair negotiation, walking, ADL and treatment success showed acceptable discriminative values (Area Under the Curve (AUC) 0.68 – 0.74). The algorithms for residual complaints regarding kneeling, squatting, pain and satisfaction showed less favourable results (AUC 0.58 – 0.64). The calibration curves showed adequate calibration for most of the models.

Conclusion

A considerable proportion of patients have residual complaints after TKA. The present study showed that demographic and PROMs data collected in the LROI can be used to predict the probability of residual symptoms after TKA. The models developed in the present study predict the chance of residual symptoms for an individual patient on 10 specific items concerning treatment success, functional outcome and pain relief. This prediction can be useful for individualized expectation management in patients planned for TKA.

chapter 9

Development of preoperative prediction models for pain and functional outcome after total knee arthroplasty using The Dutch Arthroplasty Register (LROI) data

J. J. Tolk, J.H. Waarsing, R.P.A. Janssen, L.N. van Steenbergen, S.M.A. Bierma-Zeinstra, M.Reijman.

Journal of Arthroplasty. 2020 Mar;35(3):690-698.e2.

Introduction

The rate of satisfaction after total knee arthroplasty (TKA) is consistently reported around 80%, leaving 1 in 5 patients unsatisfied to some extent after their knee surgery.[1] One of the main determinants of treatment satisfaction is the fulfilment of preoperative expectations.[1–3] In this light, expectation management in TKA patients resulting in more realistic expectations, is thought to be advantageous to achieve optimal patient satisfaction.[2,4]

Individualized education about postoperative outcome should lead to more realistic patient expectations.[4–6] Previous research has identified specific expectations of TKA patients that are considered most important,[7,8] and expectations that are most often not fulfilled in TKA patients.[9] Expectations from both these categories are useful to address in pre-operative education, and include expectations on pain-relief and improvement in walking, stair negotiation, performance of daily activities, change in position, kneeling and squatting. [8,9]

It has been shown that a useful prediction on postoperative outcome can be made based on demographic factors, pre-operative pain scores and patient-reported outcome measure (PROM) scores.[10–12] However, existing outcome prediction tools mainly focus on identifying patients at risk of not benefiting from TKA in general.[13–15] Specific information on pain and functional outcome to guide pre-operative expectation management is not provided by these tools.[13–15] Furthermore, a patient might improve in general, but could experience residual complaints on some specific items which might not be distinguished by these tools. Therefore, for use in an expectation management intervention, there is a need for an outcome prediction model that provides specific information on the probability of residual pain and functional limitations after TKA. [10–12]

Multiple patient factors that are obtained from the Dutch Arthroplasty Register (Landelijke Registratie Orthopedische Implantaten [LROI]) can potentially be utilized to estimate the most likely outcome on pain and functional outcome for an individual patient.[16] If this data can be used to make a prediction of the treatment result, the data would not only be useful for measurement of quality of care on a group level, but could be of direct value for the individual patient.

Therefore, the aim of the present study was to create and validate models that predict the chance of having residual symptoms on 10 specific outcome parameters at 12-month follow-up for individual patients undergoing primary TKA for knee osteoarthritis (OA). Since there is a known difference between men and women in functional outcome after TKA, the development of the prediction models was stratified by gender.[17,18]

Material and Methods

Study setting and data collection

Data included in this study was obtained from the LROI. The LROI is a Dutch nationwide registry that collects data on all joint arthroplasties. All hospitals in the Netherlands participate in this registry, that was founded by the Dutch Orthopaedic Association in 2007. The registry had a completeness rate of 99% for primary TKAs in 2016.[19]

Patients

Data was extracted from the registry on patients with a primary TKA for osteoarthritis, who had pre-operative and 12-month follow-up PROMs registered in the period January 2015 until July 2017, (n=7071). Only data of hospitals that provided PROM data on at least 25 patients with preoperative and postoperative PROMs questionnaires was included (53 hospitals).[20]

For assessment of generalizability, data of patients that met the inclusion criteria during the study period, but did not have PROMs registered were extracted (n=31022). A comparison for patient characteristics of the groups with and without PROMs registered was made using a chi-square test for ordinal parameters and student's t-test for continuous variables.

Data collection

Patient characteristics that are included in the LROI are age, sex, operation side, general health using the American Society of Anesthesiologists (ASA) classification system (dichotomized for regression analysis purposes; ASA I-II vs ASA III-IV),[21] body mass index (BMI) in kg/m^2, smoking (yes/no), previous operation on the affected joint (yes/no) and orthopaedic vitality using the Charnley score (dichotomized for regression analysis purposes; Charnley A vs B & C).[22] Surgical variables such as type of procedure and type of implant are registered.

Table 1. Thresholds for residual symptoms

Factor to be predicted	Corresponding question	Cut-off value for non-response
1. Residual pain at rest	NRS pain rest	>3
2. Residual pain during activity	NRS pain activity	>3
3. Rising from a chair	KOOS-PS question 3 'Indicate the degree of difficulty you have experienced in the last week due to your knee problem: Rising from sitting.'	Moderate or higher (score 2-4)
4. Stair negotiation	OKS question 12 'Could you walk down a flight of stairs?'	With moderate difficulty or worse (score 1-3)
5. Walking	OKS question 4 'For how long are you able to walk before the pain in your knee becomes severe? (With or without a stick)'	5 - 15 minutes or less (score 1-3)
6. Activities of daily living	OKS question 9 'How much has pain from your knee interfered with your usual work? (including housework)'	Moderately or more (score 1-3)
7. Kneeling	KOOS-PS question 6 'Indicate the degree of difficulty you have experienced in the last week due to your knee problem: Kneeling'	Moderate or higher (score 2-4)
8. Squatting	KOOS-PS question 7 'Indicate the degree of difficulty you have experienced in the last week due to your knee problem: squatting'	Moderate or higher (score 2-4)
9. Dissatisfaction	NRS satisfaction	<8
10. No treatment success	OKS total score	<32.5

NRS; Numerical Rating Scale. KOOS-PS; Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form. OKS; Oxford Knee Score.

Patient-Reported Outcome Measures

For primary TKA patients, the Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form (KOOS-PS), Oxford Knee Score (OKS), EuroQol 5D-3L and Numerical Rating Scale (NRS) on pain and satisfaction are obtained pre-operatively, at 6- and 12-months follow-up. The individual questions and composite scores of each PROM were available for analysis.

The KOOS-PS Dutch version assesses physical function on 7 different activities. These are scored on a 5-point Likert scale ranging from none to extreme difficulty. A normalized score can be calculated ranging from 0 indicating no symptoms and 100 indicating extreme symptoms.[23] KOOS-PS has shown to be reliable, valid and responsive to change in knee OA patients.[23–26]

The Dutch version OKS is a PROM on pain and function after TKA.[27] Twelve items are scored on a 5-point Likert scale, leading to a total score ranging from 0 to 48. Lower scores indicate more symptoms.[27] It has good measurement properties in knee OA patients.[26,27] The cut-off value for treatment success after TKA is set at >32.5 points on the OKS total score.[28]

The Dutch version of the EuroQol 5D-3L is a PROM on health-related quality of life (HRQoL).[29] The score consists of 5 questions (EQ-5D index) and the EQ visual analogue scale (EQ-VAS). The EQ-VAS records the respondent's self-rated health on a vertical, visual analogue scale where the endpoints are labelled 'Best imaginable health state' and 'Worst imaginable health state'. The 5 questions are scored on a 3-point Likert scale, from which the EQ-5D index can be calculated. The outcome scores range from -0.333 to 1.0, where 1.0 represents perfect HRQoL. This questionnaire has good reliability and validity in knee OA patients and has been validated for the Dutch population.[30,31]

Pain scores were measured using an NRS for pain during activity and in rest. A score of 0 represented 'no pain' and a score of 10 represented 'worst imaginable pain'. NRS pain values of ≤3 correspond with none or mild pain, and NRS pain score >3 represents moderate to severe pain.[32] The NRS has good reliability and responsiveness.[24]

Satisfaction with treatment results at 12-month follow-up is scored on an NRS scale ranging from 0 (not satisfied at all) to 10 (best possible satisfaction). 'Very satisfied with treatment result' is defined as an NRS satisfaction score of ³8.

Outcome of interest

The developed prediction tools provide a probability of persistent complaints on items that have been identified as important and/or most often unfulfilled in TKA patients.[7,9,33] The items addressed are residual symptoms concerning pain at rest and during activity, sit-to-stand movement, stair negotiation, walking, performance of activities of daily living, kneeling and squatting. Furthermore, satisfaction and treatment success in general were included as outcome parameters of interest in the predictive model development. For each of these factors, a corresponding question from the available PROMs in the LROI data set was identified and a threshold for remaining symptoms was chosen (Table 1).

Statistical analysis

The LROI data was randomly divided into two sets using SPSS Statistics version 24.0 (IBM Corporation). Part one of the dataset (70% of patients) was used for building the prediction models, and the remaining 30% of the dataset for validation of the models.

Model development

Gender differences play an important role in outcome after TKA, therefore the prediction model development was performed for men and women separately.[17,18] Patient characteristics and baseline PROM scores (questions and total scores) were used as candidate predictors. Categorical variables were presented using frequency and percentage, and continuous variables as mean with standard deviation (SD). Candidate variables with more than 25% missing data were excluded. Patterns of missingness were investigated to assess the presence of a non-random element to the missing data. For the remaining variables, multiple imputations were performed to estimate the missing values, resulting in 5 imputed data sets.[34]

A logistic regression analysis was performed for each of the 10 dependent variables listed in Table 1, for men and women separately. Potential predictors were identified in the univariable analyses, with significance set at $p < 0.15$. Potential predictors identified in the univariable analysis were entered into the multivariable model. This resulted in a predictive algorithm for each outcome variable of interest.

We assessed the predictive performance of these algorithms by examining measures of calibration and discrimination. Discrimination is the ability of the prediction model to distinguish between patients that have residual complaints after TKA from patients that do not have complaints. This was assessed by calculating the area under the receiver operating characteristic curve statistic (AUC) and Nagelkerke's R^2 as a measure of explained variation. Calibration considers the agreement between the predicted and the actual outcome. This was assessed using calibration plots, in which patients were classified by tenths of the predicted risk, augmented by a locally estimated scatterplot smoothed (loess) line over the entire predicted probability range.[35] Predictions of a perfect model should lie on the 45-degree line for agreement with the actual outcome.

All statistical analyses were performed using SPSS statistics version 24.0 (IBM Corporation). This study was conducted and reported in line with the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines.[36]

Results

Between January 2015 and July 2017, 38093 patients were registered in the LROI with a primary TKA for knee osteoarthritis. Of these patients, 7071 had had pre-operative and 12-month follow-up PROMs available (flowchart figure 1). There were no candidate predictors with >25% of missing data in this cohort, and there was no non-random element to the missing data. Therefore, multiple imputation could be performed as planned.[34] When compared to the group of patients that did not have PROMs available, the cohort of patients that did have PROMs available was slightly younger (68.4 (+/-8.5) vs 68.8 (+/-9.1) years,

p<0.001), more often male (37.3% vs 35.8%, p=0.015), more often smoker (8.3% vs 9.5%, p=0.002) and had a marginally lower BMI (29.6 (+/-4.8) vs. 29.8 (+/-5.1) kg/m^2, p=0.001). Full characteristics of both groups can be found in supplementary file 1.

The 7071 patients that were available for development of the prediction model were randomly divided into two groups. A training cohort (n=4951) and a cohort for internal validation of the prediction models (n=2120). Patient characteristics of these groups did not show important differences (Table 2). The incidence of residual complaints at 1-year follow-up is presented in Table 3. Of the outcome variables under study, residual complaints on kneeling (female 72% /male 59%) and squatting (female 71% / male 56%) were reported most frequently. Least residual complaints were scored for walking (female16% / male 12%) and pain at rest (female18% / male 14%).

The significant factors identified in the univariate analyses were all entered into the multivariable regression models. The majority of the individual PROM questions and PROM total scores were included in the different models. Other patient characteristics such as Charnley score, side affected, previous surgery to the affected joint and smoking status showed to be significant univariable predictors in only a few models. In almost all models the factor that showed the highest predictive value was the pre-operative PROM question corresponding with the factor that was predicted. For example, in the model predicting residual complaints when rising from a chair for female patients, KOOS-PS question 6 (the question addressing problems with sit-to-stand movement) showed to be the strongest

Table 2. Patient characteristics and baseline scores of training and test cohort.

	Training cohort				Test cohort			
	Female (n=3120)		Male (n=1831)		Female (n=1315)		Male (n=805)	
Age, years	**68.8**	**(8.6)**	**68.1**	**(8.1)**	**68.8**	**(8.6)**	**67.8**	**(8.2)**
ASA classification, n (%)								
I	380	(12.2)	285	(15.6)	160	(12.2)	132	(16.4)
II	2213	(70.9)	1231	(67.2)	934	(71.0)	531	(66.0)
III-IV	527	(16.9)	315	(17.2)	221	(16.8)	142	(17.6)
Smoking, yes n (%)	219	(7.0)	187	(10.2)	91	(6.9)	87	(10.8)
Charnley score, n (%)								
A	1328	(42.6)	878	(48.0)	551	(41.9)	395	(49.1)
B1	1024	(32.8)	585	(31.9)	435	(33.1)	267	(33.2)
B2	663	(21.3)	320	(17.5)	284	(21.6)	125	(15.5)
C	105	(3.4)	48	(2.6)	45	(0.3)	17	(0.2)
BMI, kg/m^2	30.0	(5.2)	28.9	(4.0)	30.1	(5.4)	28.9	(4.1)
Side affected, right n (%)	1677	(53.8)	948	(51.8)	700	(53.2)	426	(52.9)
Previous surgery on affected joint, yes n (%)	852	(27.3)	787	(43.0)	369	(28.1)	351	(43.6)
EQ-VAS	67.1	(19.0)	72.0	(18.7)	67.5	(19.2)	71.5	(18.0)
EQ-index	0.57	(0.27)	0.66	(0.23)	0.58	(0.27)	0.65	(0.23)
KOOS-PS total	52.9	(15.1)	47.5	(14.4)	52.8	(14.7)	47.8	(13.8)
OKS total	22.3	(6.8)	25.4	(6.7)	22.3	(7.1)	25.2	(7.1)
NRS Pain activity	7.4	(1.9)	7.0	(2.0)	7.4	(1.8)	7.0	(2.1)
NRS Pain rest	5.3	(2.6)	4.6	(2.6)	5.3	(2.5)	4.7	(2.6)

Data presented as mean (standard deviation) or otherwise as mentioned. ASA classification; American Society of Anesthesiologists (ASA) classification system. BMI; Body Mass Index. EQ; EuroQol. VAS; visual analogue scale. KOOS-PS; Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form. OKS; Oxford Knee Score. NRS; Numerical Rating Scale.

predictor with an odds ratio of 1.61 (95% CI 1.374 to 1.882).

The full prediction models are presented as clinical calculators in supplementary file 2 for female patients and supplementary file 3 for male patients (available online only). The calculator presents an individual patients' chance of residual symptoms concerning pain at rest and during activity, sit-to-stand movement, stair negotiation, walking, performance of activities of daily living, kneeling and squatting and the chance of dissatisfaction and no overall treatment success.

Table 3. Frequency of residual complaints.

Dependent variable	Sex	Training cohort (♀ n=3120, ♂ n=1831)	Validation cohort (♀ n=1315, ♂ n=805)
Residual pain rest (n [%])	♀	568 (18.2)	230 (17.5)
	♂	443 (14.2)	115 (14.3)
Residual pain activity (n [%])	♀	886 (28.4)	368 (28.0)
	♂	434 (23.7)	188 (23.4)
Rising from a chair (n [%])	♀	721 (23.1)	301 (22.9)
	♂	361 (19.7)	158 (19.6)
Stair negotiation (n [%])	♀	876 (27.8)	359 (27.3)
	♂	297 (16.2)	140 (17.4)
Walking (n [%])	♀	493 (15.8)	199 (15.1)
	♂	220 (12.0)	93 (11.6)
Activities of daily living (n [%])	♀	761 (24.4)	330 (25.1)
	♂	342 (18.7)	161 (20.0)
Kneeling (n [%])	♀	2237 (71.7)	940 (71.5)
	♂	1086 (59.3)	483 (60.0)
Squatting (n [%])	♀	2228 (71.4)	948 (72.1)
	♂	1038 (56.7)	459 (57.0)
Dissatisfaction (n [%])	♀	933 (29.9)	396 (30.1)
	♂	450 (24.6)	197 (24.5)
No treatment success (n [%])	♀	658 (21.1)	284 (21.6)
	♂	258 (14.1)	116 (14.4)

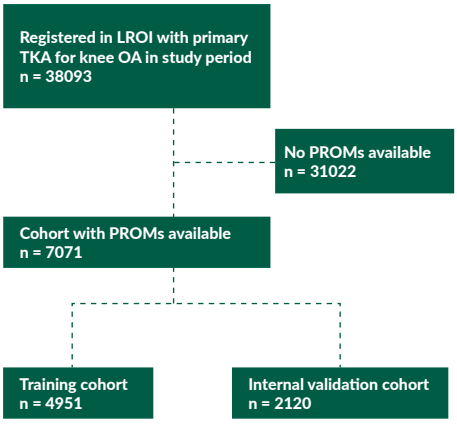


Figure 1. *Flowchart for inclusion and exclusion of patients.*

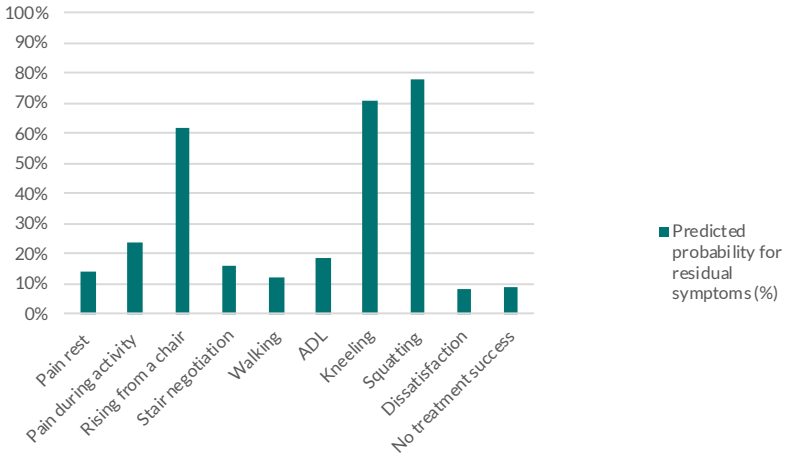


Figure 1. *Example of output provided by the predictive model. The individual chances of residual symptoms for a patient on each of the 10 items presented are calculated based on patient characteristics, baseline EuroQol-5D scores, Oxford Knee Score, and Knee injury and Osteoarthritis Outcome Score - Physical Function Short Form scores.*

An example of the output provided by the prediction model is presented in figure 2. The performance of these models in terms of discrimination are presented in Table 4 for both the training and validation cohort. The predictive algorithms for residual symptoms concerning rising from a chair, stair negotiation, walking, activities of daily living and treatment success showed acceptable discriminative values (AUC 0.68 – 0.74) and explained fraction of variance (Nagelkerke R^2 0.13-0.21). The prediction models for residual complaints regarding kneeling, squatting, pain and satisfaction showed the least favourable results on discrimination (AUC 0.58 – 0.64) and explained variance (Nagelkerke R^2 0.04-0.11). The calibration of the models in the validation cohort is presented using calibration curves. Figure 3 shows the performance of the models for women and Figure 4 for men.

Discussion

We have developed and internally validated prediction models for residual symptoms 12 months after TKA. The models predict the chance of residual symptoms on a set of outcome parameters that are considered most important in pre-operative expectation management. The output of the models presents an individual patients' chance of residual complaints (figure 2). These models generally showed fair discrimination and good calibration.
The present study shows that a considerable proportion of patients has residual complaints after TKA. Residual complaints on kneeling and squatting were reported most frequently, while the least residual complaints were reported for walking and pain at rest. Previously, TKA patients have been reported to still experience substantial functional impairment compared with their age-matched peers, especially in biomechanically demanding activities.[9,37] The results of the present study emphasize the fact that residual complaints are common and this should be considered in pre-operative decision making and expectation management.

In the study design we chose to develop separate prediction models for male and female patients. Residual complaints showed to be considerably more frequent in women over the whole range of outcome parameters that were analysed. This confirms the conclusion of previous reports in this regard,[17] and supports the choice to develop separate prediction models for male and female patients in the present study. The predictive performance of the models did not show important differences between sexes.

The discriminative performance of most prediction models developed in the present study can be considered acceptable, with AUC values in the internal validation cohort of around 0.70.[38] The prediction models for residual complaints regarding kneeling, squatting, pain and dissatisfaction with treatment result showed the least favourable results. Because of the study design, prognostic variables included were limited to the variables registered in the LROI. As a result, not all pre-operative factors that are known to influence outcome after TKA, such as pre-operative expectations and radiological OA severity, could be included as predictors for the model.[10,39] Therefore, their influence on outcome and patient satisfaction could not be incorporated by the models, which might have limited performance of the models. At the same time, by including only predictors that patients currently already provide, we made sure the models can easily be integrated in clinical practice.

For most outcome variables adequate calibration curves are reported. For predicted probabilities of residual complaints exceeding 0.5, the loess lines tend to deviate from the ideal line. This is probably due to the low numbers of patients in this category. Therefore, predicted probabilities of residual complaint exceeding 50% have to be interpreted with some caution. Considering residual complaints on kneeling and squatting the distribution of predicted probabilities was reversed in comparison to all other categories. This is in line with the relatively high prevalence of residual complaints; 57% for male and 72% for female patients. For kneeling and squatting prediction for probabilities of 40% and higher seem adequately calibrated.

To the authors' knowledge, this is the first time such a comprehensive set of prediction models on treatment result after TKA is reported. Previous efforts have been made to predict the treatment result of TKA patients.[11,13–15] These studies mainly aimed to predict overall response or treatment success in general.[11,13–15] For this purpose, Dowsey et al. created a prognostic nomogram predicting probabilities of nonresponse to TKA.[10] A limited set of predictors was derived, and internal validation showed acceptable calibration and discrimination (c-statistic 0.74) quite similar to the performance of our models on treatment success.[10] Unfortunately, these favourable results were only partly supported on external validation of this predictive nomogram by Riddle et al.[13] Especially calibration showed poor agreement between actual vs predicted probabilities of nonresponse.[13] Therefore, the model developed by Riddle et al. does not seem applicable beyond the population in which it was developed.[10,13] We intend to use the models presented in the present study in a very comparable population as in which they were developed. Therefore, we do not expect similar calibration problems.

For effective expectation management, a prediction tool should ideally provide specific information on the most likely outcome on pain and function for an individual patient.[6] Pua et al. constructed a model on risk for walking limitations after TKA.[11] In contrast to the

models described in the previous paragraph that predict outcome in general, the model by Pua et al. is the only predictive model for a specific functional outcome parameter available in literature. The model is based on a predictor set that has limited overlap with the predictors used in the present study, and used data on postoperative recovery in addition to pre-operative measurements. The predictive performance of their nomogram seems to be quite comparable to our model on walking limitations, with a reported c-statistic of 0.71.[11]

The candidate predictors for the prediction models constructed in the present study are part of routinely collected data. Since this information is always obtained for patients in the Dutch system, it was not deemed necessary to reduce the number of predictors to obtain the smallest predictor set possible.[36] Further reduction would only lead to loss of predictive power without an increase in usability or reduction of the burden for patients. The data entered in routinely obtained questionnaires can directly be transformed into an individualized prediction for probability of residual complaints (example output in figure 2). These predictions are a good basis for improved pre-operative expectation management.

A limitation of the present study is that only 19% of the complete cohort of TKA patients had completed PROMs at baseline and 12 months follow up. Therefore, only this subset of the total cohort could be included for development of the models. This is probably partly caused by the knee PROM follow-up program was only introduced in 2015 in the Netherlands and implementation has just recently started in some clinics. To avoid selection bias in hospitals that included only a very few patients, hospitals with <25 patients registered were excluded from the analysis. The group included for modelling statistics and the group that could not be included showed significant differences for some patient characteristics, but these differences were very small and therefore probably not clinically relevant. For this reason, the results of the present study can be considered generalizable to the whole Dutch population of patients undergoing primary TKA for knee OA. Patient characteristics, methods of measurement and healthcare systems might differ across populations, and can potentially influence predictive performance.[13] These models are specific for the Dutch healthcare setting. Generalization beyond this population would therefore warrant external validation and possibly recalibration. Nevertheless, the concept of outcome prediction based on routinely gathered patient characteristics and PROMs is likely to be applicable beyond the Dutch healthcare setting.

Another limitation is that although we showed that a useful prediction can be made for the probability of residual symptoms after TKA, the predictions cannot be used as explicit indication criteria for TKA. A cut off value for appropriateness of TKA is not provided and the level of discrimination would not justify such strong conclusions. Furthermore, the indication for TKA has a subjective nature where each patient has to consider the risks and benefits against their own values.[16] In the authors' opinion, the prediction results are useful for identifying patients at risk for residual complaints, and individualizing expectation management in this regard.

Conclusion

A considerable proportion of patients have residual complaints after TKA. The present study showed that demographic and PROMs data collected in the LROI can be used to predict the probability of residual symptoms after TKA. The models developed in the present study predict the chance of residual symptoms for an individual patient on 10 specific items concerning functional outcome and pain relief. This prediction can be useful for individualized expectation management in patients planned for TKA.

Funding

This study was supported by a grant from the van Rens Funds, and conducted with data of the Dutch Arthroplasty Registry.

Table 4. Performance of predictive models – discrimination.						
		Training data set			Validation data set	
Dependent variable	Sex	AUC (95%CI)		Nagelkerke R2	AUC (95%CI)	
Pain rest	♀	0.69	(0.66-0.72)	0.11	0.65	(0.61-0.69)
	♂	0.68	(0.64-0.72)	0.09	0.62	(0.57-0.67)
Pain activity	♀	0.65	(0.63-0.67)	0.08	0.63	(0.60-0.66)
	♂	0.68	(0.65-0.71)	0.11	0.64	(0.62-0.66)
Rising from a chair	♀	0.76	(0.74-0.78)	0.21	0.73	(0.70-0.76)
	♂	0.73	(0.71-0.75)	0.16	0.70	(0.66-0.74)
Stair negotiation	♀	0.72	(0.70-0.74)	0.17	0.71	(0.69-0.73)
	♂	0.72	(0.69-0.75)	0.13	0.69	(0.66-0.72)
Walking	♀	0.74	(0.72-0.76)	0.16	0.72	(0.69-0.75)
	♂	0.71	(0.69-0.73)	0.13	0.69	(0.67-0.71)
Activities of daily living	♀	0.70	(0.67-0.73)	0.13	0.67	(0.64-0.70)
	♂	0.72	(0.70-0.74)	0.14	0.67	(0.64-0.70)
Kneeling	♀	0.65	(0.63-0.67)	0.08	0.64	(0.60-0.68)
	♂	0.65	(0.62-0.68)	0.09	0.62	(0.58-0.66)
Squatting	♀	0.64	(0.62-0.66)	0.07	0.63	(0.60-0.66)
	♂	0.63	(0.61-0.65)	0.08	0.60	(0.57-0.63)
Dissatisfaction	♀	0.61	(0.59-0.64)	0.04	0.58	(0.55-0.61)
	♂	0.63	(0.60-0.66)	0.06	0.59	(0.55-0.63)
No treatment success	♀	0.73	(0.70-0.76)	0.16	0.70	(0.67-0.73)
	♂	0.73	(0.71-0.75)	0.15	0.69	(0.67-0.71)

AUC; Area Under the Curve. CI; confidence interval.

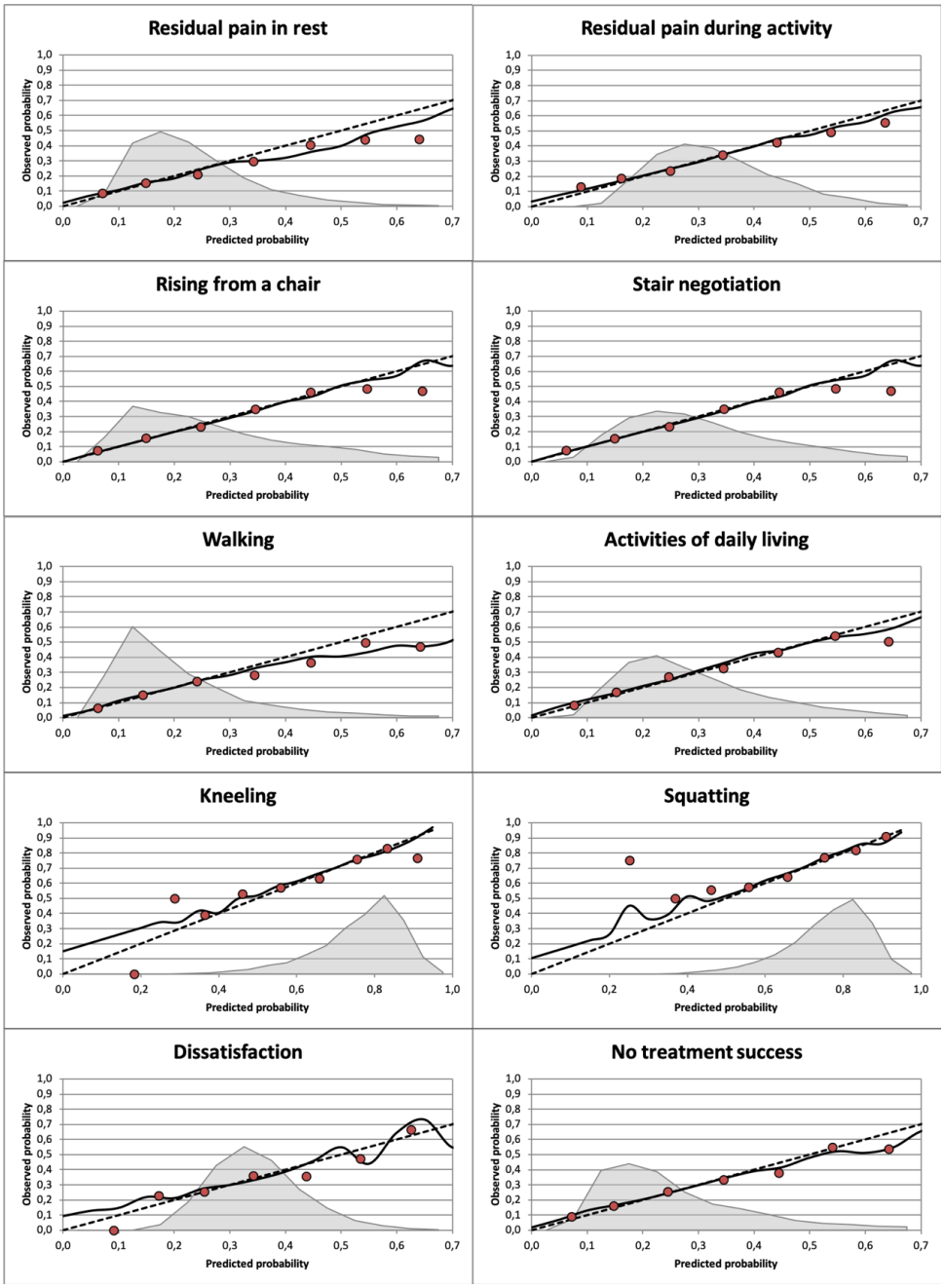


Figure 3. *Calibration curves for female patients. Predicted probability of residual symptoms on each specific outcome parameter is given on the x-axis, and the observed probability of residual symptoms is given on the y-axis. The dashed diagonal line represents perfect agreement between the predicted and actual probability of nonresponse. Tenths of the predicted risk are presented as reds dots, and augmented by a smoothed (loess) line over the entire predicted probability range (continuous line). In grey above the X-axis the distribution of the predicted probabilities.*

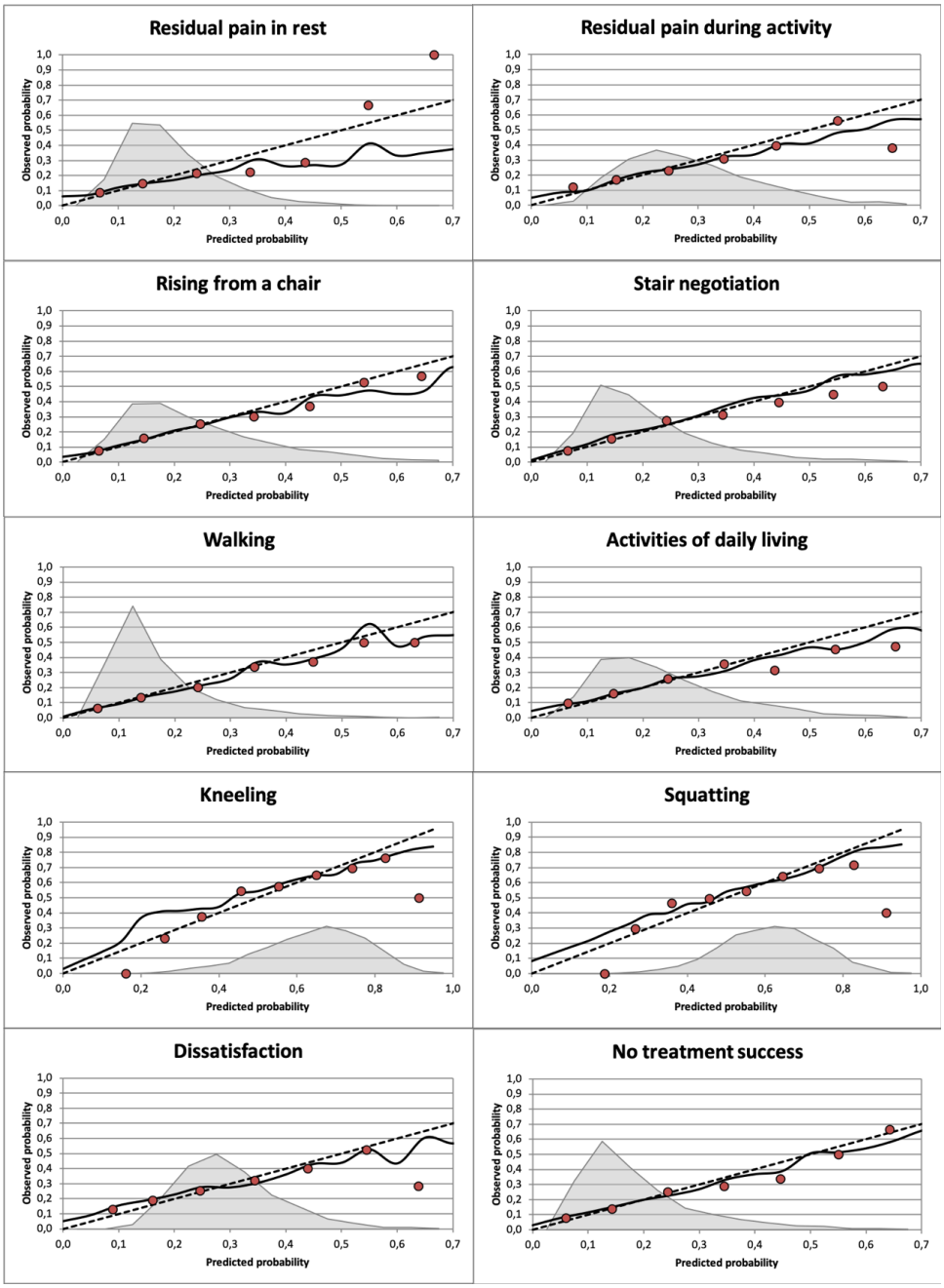


Figure 4. *Calibration curves for male patients. Predicted probability of residual symptoms on each specific outcome parameter is given on the x-axis, and the observed probability of residual symptoms is given on the y-axis. The dashed diagonal line represents perfect agreement between the predicted and actual probability of nonresponse. Tenths of the predicted risk are presented as reds dots, and augmented by a smoothed (loess) line over the entire predicted probability range (continuous line). In grey above the X-axis the distribution of the predicted probabilities.*

Supplementary file 1. Patients with PROM data vs patient without PROM data registered in LROI.			
	Cohort with PROMs available (n=7071)	**Cohort with no PROMs available (n=31022)**	**p-value***
Sex female, n (%)	4435 (62.7)	19961 (64.2)	p=0.015
Age	68.4 (8.5)	68.8 (9.1)	p<0.001
ASA classification, n (%) I II III-IV	 962 (13.6) 4886 (69.1) 1223 (17.3)	 4219 (13.6) 21591 (69.6) 5212 (16.8)	p=0.615
Smoking, yes n (%)	587 (8.3)	2947 (9.5)	p=0.002
Charnley score, n (%) A B1 B2 C	 3161 (44.7) 2305 (32.6) 1393 (19.7) 212 (3.0)	 13246 (42.7) 10547 (34.0) 6356 (21.1) 838 (2.7)	p<0.001
BMI, kg/m^2	29.6 (4.8)	29.8 (5.1)	p=0.001
Side affected, right n (%)	3733 (52.8)	16256 (52.4)	p=0.809
Previous surgery on affected joint, yes n (%)	2376 (33.6)	10389 (33.5)	p=0.562
Independent-Samples T-test or Chi-Square depending on the type of variable.			

References

1. Noble PC, Conditt M a, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452(452):35-43. doi:10.1097/01.blo.0000238825.63648.1e
2. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: Who is satisfied and who is not? *Clin Orthop Relat Res*. 2010;468(1):57-63. doi:10.1007/s11999-009-1119-9
3. Hamilton DF, Lane J V, Gaston P, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open*. 2013;3(4):1-8. doi:10.1136/bmjopen-2012-002525
4. Tolk JJ, Janssen RPA, Haanstra TM, Bierma-zeinstra SMA, Reijman M. The EKSPECT study : the influence of Expectation modification in Knee arthroplasty on Satisfaction of PatiEnts : study protocol for a randomized Controlled Trial. *Trials*. 2018;19(437).
5. Mancuso C, Graziano S, Briskie L, et al. Randomized trials to modify patients' preoperative expectations of hip and knee arthroplasties. *Clin Orthop Relat Res*. 2008;466(2):424-431. doi:10.1007/s11999-007-0052-z
6. Judge A, Arden NK, Cooper C, et al. Predictors of outcomes of total knee replacement surgery. *Rheumatology (Oxford)*. 2012;51(10):1804-1813. doi:10.1093/rheumatology/kes075
7. Mancuso CA, Sculco TP, Wickiewicz TL, Jones EC, Robbins L, Warren RF. Patients' expectations of knee surgery. *J Bone Jt Surg Am*. 2001;83-A(7):1005-1012.
8. Lange T, Schmitt J, Kopkow C, Rataj E, Günther K-P, Lützner J. What Do Patients Expect From Total Knee Arthroplasty? A Delphi Consensus Study on Patient Treatment Goals. *J Arthroplasty*. 2017:1-8. doi:10.1016/j.arth.2017.01.053
9. Tilbury C, Haanstra TM, Leichtenberg CS, et al. Unfulfilled Expectations After Total Hip and Knee Arthroplasty Surgery: There Is a Need for Better Preoperative Patient Information and Education. *J Arthroplasty*. 2016;31(10):2139-2145. doi:10.1016/j.arth.2016.02.061
10. Dowsey MM, Spelman T, Choong PFM. Development of a Prognostic Nomogram for Predicting the Probability of Nonresponse to Total Knee Arthroplasty 1 Year After Surgery. *J Arthroplasty*. 2015;31:1654-1660. doi:10.1016/j.arth.2016.02.003
11. Pua YH, Seah FJT, Clark RA, Poon CLL, Tan JWM, Chong HC. Development of a prediction model to estimate the risk of walking limitations in patients with total knee arthroplasty. *J Rheumatol*. 2016;43(2):419-426. doi:10.3899/jrheum.150724
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;31:1-9. doi:10.1016/j.arth.2016.06.004
13. 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. doi:10.1016/j.arth.2016.11.007
14. Barlow T, Dunbar M, Sprowson A, Parsons N, Griffin D. Development of an outcome prediction tool for patients considering a total knee replacement--the Knee Outcome Prediction Study (KOPS). *BMC Musculoskelet Disord*. 2014;15(1):451. doi:10.1186/1471-2474-15-451
15. Lungu E, Desmeules F, Dionne CE, Belzile EL, Vendittoli P-A. Prediction of poor outcomes six months following total knee arthroplasty in patients awaiting surgery. *BMC Musculoskelet Disord*. 2014;15:299. doi:10.1186/1471-2474-15-299
16. Berliner JL, Brodke DJ, Chan V, SooHoo NF, Bozic KJ. Can Preoperative Patient-reported Outcome Measures Be Used to Predict Meaningful Improvement in Function After TKA? *Clin Orthop Relat Res*. 2016;475(1):1-9. doi:10.1007/s11999-016-4770-y

17. Cherian J, O'Connor M, Robinson K, Jauregui J, Adleberg J, Mont M. A Prospective, Longitudinal Study of Outcomes Following Total Knee Arthroplasty Stratified by Gender. *J Arthroplasty*. 2015;30:1372-1377. doi:10.1055/s-0034-1388658
18. Volkmann ER, FitzGerald JD. Reducing gender disparities in post-total knee arthroplasty expectations through a decision aid. *BMC Musculoskelet Disord*. 2015;16:16. doi:10.1186/s12891-015-0473-x
19. Dutch Arthroplasty Register (LROI). Report on data quality coverage and completeness. http://www.lroi-rapportage.nl/data-quality-coverage-and-completeness. Published 2016.
20. Peters RM, van Beers LWAH, van Steenbergen LN, et al. Similar Superior Patient-Reported Outcome Measures for Anterior and Posterolateral Approaches After Total Hip Arthroplasty: Postoperative Patient-Reported Outcome Measure Improvement After 3 months in 12,774 Primary Total Hip Arthroplasties Using the Ante. *J Arthroplasty*. 2018;33(6):1786-1793. doi:10.1016/j.arth.2018.01.055
21. Fitz-Henry J. The ASA classification and peri-operative risk. *Ann R Coll Surg Engl*. 2011;93(3):185–187.
22. Charnley J. The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. *J Bone Jt Surgery, Br …*. 1972;54(1):61-76.
23. Perruccio A V, Lohmander LS, 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. *Osteoarthr Cart*. 2008;16:542-550. doi:10.1016/j.joca.2007.12.014
24. Ruyssen-Witrand A, Fernandez-Lopez CJ, Gossec L, Anract P, Courpied JP, Dougados M. Psychometric properties of the OARSI/OMERACT osteoarthritis pain and functional impairment scales: ICOAP, KOOS-PS and HOOS-PS. *Clin Exp Rheumatol*. 2011;29(2):231-237.
25. de Groot IB, Favejee MM, Reijman M, Verhaar J a N, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. *Health Qual Life Outcomes*. 2008;6:16. doi:10.1186/1477-7525-6-16
26. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Outcome. *Arthritis Care Res (Hoboken)*. 2011;63 Suppl 1(November):S208-28. doi:10.1002/acr.20632
27. Haverkamp D, Breugem SJM, 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 Orthop*. 2005;76(3):347-352. doi:10.1080/00016470510030814
28. Hamilton DF, Loth FL, MacDonald DJ, et al. Treatment Success Following Joint Arthroplasty: Defining Thresholds for the Oxford Hip and Knee Scores. *J Arthroplasty*. 2018;33(8):2392-2397. doi:10.1016/j.arth.2018.03.062
29. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343.
30. Lamers L, Stalmeier P, McDonnell J, Krabbe P, van Busschbach J. Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff. *Ned Tijdschr Geneeskd*. 2005;149(28):1574-1578.
31. Conner-Spady BL, Marshall DA, Bohm E, et al. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. *Qual Life Res*. 2015;24(7):1775-1784. doi:10.1007/s11136-014-0910-6
32. Pinto PR, 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). doi:10.1371/journal.pone.0073917

33. Smith EJ, Soon V-L, Boyd A, McAllister J, Deakin A, Sarungi M. What do Scottish Patients Expect of Their Total Knee Arthroplasty? *J Arthroplasty*. 2016;31(4):786-792. doi:10.1016/j.arth.2015.10.021
34. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - A practical guide with flowcharts. *BMC Med Res Methodol*. 2017;17(1):1-10. doi:10.1186/s12874-017-0442-1
35. Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. *Stat Med*. 2014;33:517-535. doi:10.1002/sim.5941
36. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Ann Intern Med*. 2015;162(1):W1. doi:10.7326/M14-0698
37. Noble PC, Gordon MJ, Weiss JM, Reddix RN, Conditt M a, Mathis KB. Does total knee replacement restore normal knee function? *Clin Orthop Relat Res*. 2005;(431):157-165. doi:10.1097/01.blo.0000150130.03519.fb
38. Gortmaker S, Hosmer D, Lemeshow S. Applied logistic regression. *Contemp Sociol*. 1994;23:159.
39. Jain D, Nguyen L-CL, Bendich I, et al. Higher patient expectations predict higher patient-reported outcomes, but not satisfaction, in total knee arthroplasty patients: a prospective multicenter study. *J Arthroplasty*. 2017;32(9S):S166-S170. doi:10.1016/j.arth.2017.01.008

chapter 10

General discussion

This thesis focuses on patients with knee and hip osteoarthritis undergoing total knee arthroplasty (TKA) and total hip arthroplasty (THA) respectively. In the previous chapters it was discussed how outcome can be assessed, expectations of treatment result and whether expectation management can influence satisfaction after surgery.

PART 01 | Outcome assessment after arthroplasty

As stated in the introduction TKA and THA can generally be considered as successful treatments. Nevertheless, within the multiple dimensions of clinical outcome defining 'treatment success' is multi-interpretable. Traditionally orthopaedic surgeons primarily focused on the technical result of the surgical procedure and absence of complications, whereas patients are most interested in symptom relief and resumption of physical activities.[1,2] The technical and patient experienced result can't be seen independently, and with an increasingly younger and more active patient population the patient's perspective has been firmly established in outcome evaluation in current clinical care.

From the patient's perspective physical function is an important outcome parameter that should be considered in outcome assessment after knee and hip arthroplasty.[3,4] Whereas in clinical practice measurement in this domain nowadays mainly relies on patient reported outcome measures (PROMs), these instruments do have distinct disadvantages. This is why it has been advocated to use performance-based measures (PBMs) to evaluate physical function.[25,26] Part 1 of this thesis evaluated the measurement properties of a core set of performance-based measures **(chapter 2 and chapter 3)**. In this section the current state of knowledge and challenges in outcome assessment after TKA and THA will be discussed, in the light of the specific advantages and disadvantages of both PROMs and PBMs.

Patient reported outcome measures

In reporting clinical trials on hip and knee osteoarthritis a recent update of the OMERACT/OARSI criteria considers at least 5 domains as mandatory to be measured and reported: pain, physical function, quality of life, patient's global assessment of the target joint and adverse events.[3] The first four can be considered patient driven outcomes and largely coincide with the Dutch Orthopaedic Association (Nederlandse Orthopedische Vereniging [NOV]) as well as the International Consortium for Health Outcomes Measurement (ICHOM) advice on patient reported outcome measurement in joint arthroplasty.[4] The NOV and ICHOM advocate the use of specific PROMs to assess pain (NRS-pain), quality of life and global assessment (EQ-5D / EQ-VAS) and joint specific scores for assessment of physical function. To measure physical function for hip arthroplasty patients the Hip disability and Osteoarthritis Outcome Score - Physical Function Short Form (HOOS-PS) and Oxford Hip Score (OHS) are recommend and for TKA patients the Knee disability and Osteoarthritis Outcome Score - Physical Function Short Form (KOOS-PS) and Oxford Knee Score (OKS). Both the HOOSps/KOOSps and OHS/OKS scores are well-accepted outcome measures. These disease specific questionnaires aim to measure physical function in hip and knee osteoarthritis patients respectively, and are reported to have good clinimetric properties in terms of reliability, validity and responsiveness.[5–9] Nevertheless, when assessing the construct physical function using PROMs, some inherent limitations have to be acknowledged.

Important items concerning physical function are covered by the recommended questionnaires, but they are not comprehensive with regard to the full definition of activity and participation according to the ICF (International Classification of Functioning, Disability and Health).[10,11] Relevant activities might therefore be missed, especially in the view of increasing numbers of younger patients and increasing demands in terms of desired activity levels including strenuous recreational activities and sports.[12] Furthermore, ceiling effects are an issue of concern. Despite the fact that for the OHS and OKS no ceiling effects are reported, in subgroups such as patients with a high pre-operative level of functioning and younger patients, relevant numbers of patients have maximum outcome scores.[13,14] These findings limit the possibility to separate performance at the higher ends of the questionnaires scale. On a group level this can result in reporting no difference between groups when there actually is, for example in evaluating improvement in arthroplasty design or rehabilitation strategies. For individual patients this could influence decision-making, as for some patients maximum scores do not reflect full recovery on all potential activities that might be relevant to them.

To address these issues efforts have been made by adapting scores to allow for assessment of more high demand functional activities. For example, using the forgotten joint score or the high activity arthroplasty score (HAAS).[15,16] These questionnaires have proven useful for example in research settings when further detail regarding the levels of activity and participation of a patient is required.[15–17] Nevertheless, they are designed to be used alongside original PROMs, and not recommended as a stand-alone instrument.[18] In clinical practice feasibility issues and patient burden limits the amount and length of questionnaires that should be used. Moreover, increased length of questionnaires is related to lower response rates as well as lower quality of the answers.[19] Therefore adding more questionnaires to overcome ceiling effects might not be the best approach. Implementation of the Patient-Reported Outcomes Measurement Information System (PROMIS®) seems a promising solution for this problem.[10,20,21]

The PROMIS instruments have been developed using Item Response Theory and consist of item banks with a range of questions representing increasing symptoms on a single construct, for example physical functioning.[20,21] One of the main advantages is that PROMIS can be applied as a computer adaptive test (CAT). A computer algorithm selects questions based on the responses to previous items, presenting only relevant items to a patient. Using a small and flexible number of items (usually 3 to 7), the CAT can efficiently come to a valid and precise score, that is easily interpretable.[20] The Dutch-Flemish PROMIS Physical Function item bank has been appropriately evaluated with good results in physiotherapy patients.[20,22] Application in knee and hip arthroplasty patients can certainly be considered, but further clinimetric evaluation in this population, especially on responsiveness, is still necessary.[20,23]

Performance based measures

Because of the strong relationship between pain and physical function measured using PROMs using self-reported instruments,[24] it is difficult to distinguish between the effect of pain reduction and actual improvement in physical function. To avoid both ceiling effects and the effect of pain reduction concurrent analysis of physical function with PBMs is advocated to quantify the execution of a specific task associated with function.[25,26] In

chapter 2 and 3 the measurement properties of the recommended OARSI core set of PBMs were analysed. The tests showed to be highly reliable, but unfortunately the validity and responsiveness could not be confirmed.[27,28]

To assess the measurement properties of the PBMs we chose the hypothesis testing method, which is the preferred method in the absence of a gold standard.[29] An inherent limitation of this method is that you have to rely on comparator instruments that do not asses the domain under study in exactly the same way. The main driver behind the conclusion that the PBMs showed limited validity and responsiveness, was the almost completely absent relationship with patient's experienced assessment of function. The discrepancy between experienced performance of a task, and timed measures of performance has been reported before,[30] but from a clinical and patient-centred perspective this makes the PBMs under study unsuitable for clinical practice. Apparently, patients do not find their walking speed of much importance in assessing the impairment experienced in this functional domain.

To assess the performance of a task itself it might be better to assess the biomechanical quality of movement when performing the task at hand, than the speed at with it is conducted. Diminished quality of movement (eg. limping, guarding or rigidity) probably affects the patient's perception of impairment more than the time it takes. This is supported by a recent study by Biggs et al., analysing the relationship between patient reported function and biomechanical assessment of function.[31] They found that change in objective measures of the biomechanical quality of movement, were strongly correlated to changes in OKS,[31] as opposed to the virtually absent relationship between PBM's and self-reported function we and others have reported earlier.[27,28,30,32] Furthermore, thus far no relationship between PBM scores and satisfaction after TKA and THA has been established. This adds to the idea that merely timing the speed at which a task is performed or counting repetitions, do not represent very relevant measures for the assessment of physical function in clinical practice.

Use of outcome measures in clinical practice

Outcome measured using PROM's and PBM's can be considered useful for measuring differences in research setting and quality of care on a group level. For individual patients, absolute values on either PROMs or PBM's might be less important. Scores can be used to quantify an overall difference in (perceived) performance or compare a patient's complaint to other patients, but these are factors that are probably not most important to each individual patient.[33] What does matter to an individual patient is very context dependent; which specific impairments made him/her choose surgery, and affected his/her wellbeing the most. Treatment success on an individual basis should therefore be assessed in the light of a patient's individual circumstances, wishes and desires. This approach fits the value-based-healthcare approach, placing the treatment value from a patients' perspective central in clinical decision-making process.[34] To incorporate such an approach in clinical practice, it is warranted to structurally assess and discuss a patients' pre-operative expectations, and judge the individual treatment success by symptom relieve or increase in physical function on items that are actually important in a patients' life.

Addressing specific items of PROMs that are most important in a person-centred care approach is feasible on an individual basis. On group level however, this poses a problem.

The weighting of the components of PROMs scales and standardized outcome sets (eg. ICHOM) is fixed. This is generally based on average population preferences and developer standards. To add preference-sensitivity to these scores, a standard method of addressing patient sensitivity to the scales in the form of Patient Reported Importance Measures (PRIMs) is suggested.[35] These are parallel questionnaires that address a patients rating in importance of outcome components. Introduction of these measures would make it possible to not only adjust for case-mix in value-based health care, but also for structured preference-mix adjustment. Further development and validation of such PRIM adjusted PROMs is warranted.[35] Potentially this method leads to better decision making in, and evaluation of personalized value-based care.

In conclusion further research on outcome measurement in arthroplasty patients should focus on exploring the promising results of novel outcome measures that avoid floor and ceiling effects and reduce the questionnaire time burden by means of computer adaptive testing, with integration of patient preferences into the assessment of the treatment result.

Part 02 | Expectations of treatment result

The second part of this thesis addressed what can be considered realistic expectations for treatment result after TKA and analysed determinants of patients' expectations.

This thesis focused on expectations for long term recovery after knee or hip arthroplasty. This focus was mainly chosen because of the known and possibly modifiable relationship between fulfilment of these outcome expectations and postoperative satisfaction.[2,36–38] Nevertheless, expectations on long term recovery are not the only important issues for patients. Patients have expectations concerning the immediate perioperative period, rehabilitation process and speed of functional recovery.[39,40] These factors may correlate less to long term satisfaction, but are known to have high impact by decreasing pre- and postoperative anxiety, coping and self-efficacy.[40] This effect is reported to subsequently result in increased rehabilitation adherence, reduced length of stay, readmission rate and costs.[40] Therefore, although not highlighted in this thesis, achieving realistic expectations on the pre-operative and early recovery period should be considered highly important as well.

A limitation of the analysis on realistic outcome expectations was that the number of specific expectation items that were addressed throughout this thesis is limited. As basis for which expectations to address, research by Mancuso et al. was mainly used.[41] They narrowed the most important expectations for knee and hip arthroplasty patients down to 19 and 18 items respectively.[41,42] Of these items pain-relief, walking ability, stair negotiation, ability to perform daily activities and ability to change positions are rated as most important by TKA patients.[43] The specific items that were included are based on qualitative techniques (focus group discussions and individual interviews), patient and expert panel review and has resulted in valid and reliable structured surveys on expectations in knee and hip arthroplasty patients. [41,44] The items addressed can therefore be considered representative for the most important expectations of most patients. Unfortunately, more uncommon symptoms or complaints are not included, and therefore important items for individual

patients might be missed.[45,46] Earlier in this chapter the advantages of preference sensitivity in outcome measurement have been addressed. These advantages are likely to apply to the evaluation of expectations as well; not only items that are important for most patients should be addressed. Ideally an expectation management intervention should individually assess what a patient finds important, subsequently this information can be used to address these items most thoroughly. Domains that are of less or no importance to a patient can be given less emphasis. This would leave more room for a person's preferences and values, and more uncommon symptoms or complaints that might be neglected in a generic education module can be addressed. Such a person-centred method is likely to improve the effect of pre-operative expectation management,[33,47] but further research is necessary to find the optimal approach.
Furthermore, what can be considered a 'realistic' expectation can be very different for individual patients, as outcomes are diverse. The conclusion from **chapter 4 and 5** are therefore usable for providing a general prediction for what patients should expect, and the results of **chapter 6** as guidance to identify patients at risk for having too high or too low expectations. Still there remains an inherent uncertainty for what can be considered a realistic outcome expectation for an individual patient, and this uncertainty should be addressed when presenting these numbers to patients. In the education module presented in **chapter 7** and evaluated in **chapter 8**, this was recognized and to overcome this issue as much as possible individual modifying factors were specifically mentioned. Modifying factors that were addressed to predict higher outcome were younger age and better pre-operative functional status, and factors predicting worse outcome were medical and psychosocial co-morbidity, higher Body Mass Index and higher pre-operative pain severity.[48–50] Although this provides patients with some guidance on what to expect, there is considerable room for improvement in this regard. The idea for the prediction model presented in **chapter 9** was born with this issue in mind. Ideally such a model would provide an adequate outcome prediction on all items to be addressed in an expectation management module.

In **chapter 6** an analysis was performed on predictors for the height of pre-operative expectation for the treatment outcome after TKA. Female sex, higher age, higher depression score, and duration of complaints > 50 months showed to be significant predictors of lower expectations. Of these factors, only depressive symptoms are potentially modifiable. It is known that pre-operative depression scores correlate with higher pain scores and worse clinical outcome after TKA,[51,52] and expectations might be a mediating factor for this relationship. In this regard, it has been suggested that interventions designed to reduce catastrophizing and depressive symptoms may have the potential to further improve joint replacement outcomes.[51,52] In my opinion, the discussion if pre-operative treatment of depressive symptoms should be considered to improve the treatment result after arthroplasty is not of much interest. If a patient is depressed it should always be considered to help a patient towards the best treatment options when possible, as mental health impairment is strongly associated with reduced health-related quality of life frequently at levels exceeding those of physical health impairments.[53] Therefore, in depressed patients one should pay extra attention to realistic expectations, but vice versa it is even more important to be cautious on the presence of depression in patients with very low expectations.

Part 03 | Expectation management in clinical practice

In the third part of this thesis we aimed to translate the potential of improved expectation management to a clinically applicable education module. The methods of the Randomised Controlled Trial (RCT) are presented in **chapter 7**, and the results of the EKSPECT study are presented in **chapter 8**. In **chapter 9** prediction models on the chance of residual symptoms after TKA are presented. These models can be useful for individualized expectation management interventions.

Whereas in the previous parts of this thesis hip osteoarthritis patients were addressed as well, in part 3 we chose to focus solely on knee osteoarthritis patients undergoing TKA. This choice was mainly based on three reasons. First, a considerable higher satisfaction rate is reported after THA compared to TKA patients. After THA 89-97% of patients is reported to be satisfied, compared to 81-88% after TKA.[54,55] Second, the occurrence of unfulfilled expectations is less frequent in THA patients.[56] Finally, expectation management interventions studied previously, reported less change in expectations in THA patients than in TKA patients after the intervention.[57] Considering these factors, in THA patients there is less room for improvement to test the concept of improving postoperative satisfaction by enhanced pre-operative expectation management. This does not mean that a positive effect of such an intervention won't be present in THA patients. Especially as the relationship between unfulfilled expectations and dissatisfaction is reported in THA patients as well.[58] Therefore, even though the concept remains to be proven, and the effect might not be as strong as in TKA patients, it is likely that enhanced expectation management would be beneficial for THA patients too.

Aim of pre-operative expectation management

The results of the RCT in **chapter 8** gives strong evidence for the beneficial effects of improved pre-operative education module about realistic expectations for treatment result after TKA on postoperative satisfaction. The study does not completely resolve the debate on the aim of pre-operative expectation management.[68,69] In literature both advocates of increasing the height of expectations,[68] as well advocates of decreasing expectations can be found [57,69]. Considering postoperative satisfaction, it is recognized that not the height of pre-operative expectations, but the fulfilment of these expectations is highly correlated to postoperative satisfaction.[36,70] This relationship is independent from the postoperative pain, quality of life and physical function scores.[36,70] Although not yet specifically reported in **Chapter 8**, analysis of the data from this study confirmed this high correlation between expectation fulfilment and satisfaction in our population. These results can be considered advocates of aiming for expectations that are not too high, as this makes them more likely to be fulfilled. On the other hand, patients' expectations for TKA patients are recommended to be as high as possible.[68] Expectations that are too low can be potentially self-fulfilling, as because of the nocebo effect and lower patient motivation they can lead to worse treatment result.[71,72] On the opposite, higher expectations are related to better treatment adherence, anxiety reduction and beneficial coping mechanisms and thus can lead to better treatment outcome.[73,74]

We found that the expectation management intervention studied in **chapter 8** lead to a decrease in level of patient expectations. Considering that these lower expectations resulted in a higher postoperative expectation fulfilment rate, and subsequently higher postoperative satisfaction, overall lower expectations seem to be advantageous. Although on a group level this holds true, further analysis is necessary to explore this on an individual level. There are likely to be subgroups of patients for whom higher expectations would be beneficial, to improve treatment outcome.[68] Identifying these subgroups would allow to specifically address individuals that could benefit from higher expectations and address them accordingly in future expectation management interventions.

Logistics of an expectation management intervention

Some remarks can be made on the methods of the expectation module. For the RCT we chose to address the education on realistic expectations in a group session. Because not all patients attended the education module, there might be room for improvement on the method of knowledge transfer. Group communication was chosen because of the easy integration into the current clinical care, limiting the burden for the study population and increasing the likelihood for patients to be willing to participate. Besides of the chance of patients not attending, the group-based approach did limit the amount of personalization of the expectation management intervention. There was time and attention for additional questions and concerns, but only on the patient's initiative. It might be useful to address individual expectations scores (for example scored on the Hospital for Special Surgery Knee or Hip Replacement Expectations Survey), and specifically address items that are scored very low or very high and discuss the reasons for these scores, and address what would be a realistic expectation for an individual patient.[41,44]

To do so would be more time consuming, and the timing of such an intervention can be up for debate too. A possibility would be to address realistic expectations more thoroughly during an outpatient visit with the orthopaedic surgeon. It is known for a long time that very short visits are not optimal in terms of quality of care as well as patient satisfaction with the consultation.[59] Although no studies on the effect of consultation duration on realistic expectations in orthopaedic surgery are available, in other fields of medicine it has been established that longer first visits have a positive influence on satisfaction with the consultation, treatment decision making and the quality of care.[60,61] On the other hand, not only consultation time, but the quality of the conversation in terms of attitude, empathy and communication strategies plays a very important role as well.[62,63] Furthermore, lengthening consultations and providing all information regarding realistic expectations at once is not the best way to go either, especially as only a very limited part of the medical information provided is memorized by patients.[64–67] It is reported that on average more than half of the information was forgotten immediately after a medical consultation, and the information that was recalled was inaccurate approximately 50% of the time. [64–66] By subdividing information and making the education process interactive using educational apps the actual as well as perceived knowledge after education can be considerably increased.[66,67] Furthermore, an educational app that addresses realistic long term expectation might increase the number of patients reached with the information module, as the burden of having to come to the hospital for an extra visit is reduced.[66] On the other hand, the possibility of addressing individual concerns and expectation that are less common is problematic using only digital information methods. Therefore, probably a combination of

digital education followed by a group or individual expectation management session where individual concerns can be addressed seems the most promising method of education management. Further research in this regard has to show the optimal way of education, and if this would increase the reach and effect of an expectation management intervention.

Improvement of expectation management by individualized outcome prediction
Another way to improve the expectation module is to become better in predicting the most likely outcome for individual patients. In **chapter 9** we attempted to construct prediction models to do so. Pre-operative demographic factors and PROMs that are routinely gathered for the Dutch Arthroplasty Register (Landelijke Registratie Orthopedische Implantaten [LROI) were used as to create and validate models that predict residual symptoms for patients undergoing primary TKA. This resulted in 10 prediction algorithms, with variable predictive power. Whereas some performed acceptable (models for residual symptoms concerning sit-to-stand movement, stair negotiation, walking, ADL and treatment success) others performed less favourable (models for residual complaints regarding kneeling, squatting, pain and satisfaction). Although the models with the best performance can be implemented in clinical practice, efforts should be made to improve the algorithms. One approach would be to try to improve the models by adding specific predictors, that are not included in the standard set of data gathered for the LROI. Radiological osteoarthritis severity and pre-operative psychological symptoms are for example known to have considerable impact on postoperative outcome.[68,75] Further research should define if adding these factors would improve the predictive performance of the models.

Furthermore, advanced modelling strategies such as machine learning are increasingly advocated in constructing predictive models.[76,77] We used the 'traditional' method of model development using logistic regression. Machine learning approaches have attracted attention, as they have shown to yield more stable prediction because of their ability to process complex, interactive nonlinear relationships between predictors in large data sets.[76,77] Artificial-intelligence and machine-learning predictive algorithms are already finding their way into our daily life, for example by automatically driving cars and recognition of spoken language.[76] In medicine superior predictive performance over traditional approaches has been shown as well, for instance in predicting mortality in sepsis patients and unplanned transfers to the intensive care unit.[78,79] Orthopedic surgery research in this regard is still in a preliminary phase.[80] Although predictive models constructed using machine learning methods are promising, others warn for the risks of overfitting of predictions by identifying spurious correlation and limitation in the databases the models are constructed on.[77] Factors such as confounding by indication and non-random missing outcome data can result in considerable bias in machine predictions.[76,77] Therefore, models should be thoroughly validated before implementation in clinical practice is justified. Nevertheless, the potential of machine learning seems promising, and warrants further research in outcome prediction methods for arthroplasty patients. It should be recognized though, that no matter how smart the computer algorithm, it can't squeeze out information that is not there.[76] Therefore, always an inherent uncertainty regarding treatment result will be present. It will remain the doctor's role to adequately address this uncertainty in treatment decision making and pre-operative expectation management.

Future perspectives

Throughout this discussion section several aims for further research have been pointed out. In a future perspective these could come together to a comprehensive system. In such an approach pre-operative pain and function scores would be measured using efficient PROMs that allow for measuring individual patient preferences, and these scores would be used for individual outcome prediction. This prediction could in turn directly be presented to a patient for individualized expectation management.

Although a promising spot on the horizon, to come to such a comprehensive approach further research is necessary. First, to examine how the burden of measurement can be reduced without affecting the validity and interpretability of the results, and in the same time include preference sensitivity (for example using PRIMs). Validity studies on such novel outcome measures should be conducted in a Dutch healthcare setting. Second, in order to increase the performance of individualized outcome prediction, advanced modelling strategies and the effect of adding specific predictors should be attempted. Third, in the development and optimization of information transfer regarding realistic expectations, patient involvement is indispensable. Qualitative research on patient preferences and perception, and quantitative research of the effect on knowledge levels should be conducted to come to the best communication strategy. Finally, the impact of all these improvements on the reach and effect of improved expectation management interventions on postoperative satisfaction, as well as on functional outcome should be assessed.

Overall, acquiring scores and presenting individualized predictions will probably be best facilitated by ehealth solutions. Nevertheless, there will remain an important role for healthcare professionals to address a persons' specific needs, preferences and values and put the results in a personal perspective.

References:

1. Ramkumar PN, Harris JD, Noble PC. Patient-reported outcome measures after total knee arthroplasty: A systematic review. *Bone Jt Res*. 2015;4(7):120-127. doi:10.1302/2046-3758.47.2000380
2. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: Who is satisfied and who is not? *Clin Orthop Relat Res*. 2010;468(1):57-63. doi:10.1007/s11999-009-1119-9
3. Smith TO, Hawker GA, Hunter DJ, et al. The OMERACT-OARSI Core Domain Set for Measurement in Clinical Trials of Hip and/or Knee Osteoarthritis. *J Rheumatol*. 2019;46(8):981-989.
4. Patient Reported Outcome Measures - Advies Nederlandse Orthopaedische Vereniging (NOV). https://www.orthopeden.org/downloads/32/advies-proms-orthopedie.pdf. Published 2012. Accessed October 15, 2019.
5. Reiman MP, Manske RC. The assessment of function: How is it measured? A clinical perspective. *J Man Manip Ther*. 2011;19(2):91-99. doi:10.1179/10669811 1X12973307659546
6. Terwee CB, Mokkink LB, Steultjens MPM, Dekker J. Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Rheumatology (Oxford)*. 2006;45(7):890-902. doi:10.1093/rheumatology/kei267
7. Davis AM, Perruccio A V, Canizares M, et al. Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. *Osteoarthr Cartil*. 2009;17(7):843-847. doi:10.1016/j.joca.2009.01.005
8. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *Bone Joint J*. 1996;78(2):185-190. doi:10.1302/0301-620X.80B1.7859
9. Gosens T, Hoefnagels N, de Vet R, et al. The "Oxford Heup Score": the translation and validation of a questionnaire into Dutch to evaluate the results of total hip arthroplasty. *Acta Orthop*. 2005;76(2):204-211.
10. de Groot IB, Favejee MM, Reijman M, Verhaar J a N, Terwee CB. The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study. *Health Qual Life Outcomes*. 2008;6:16. doi:10.1186/1477-7525-6-16
11. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Outcome. *Arthritis Care Res (Hoboken)*. 2011;63 Suppl 1(November):S208-28. doi:10.1002/acr.20632
12. Stratford PW, Kennedy DM. A Comparison Study of KOOS-PS and KOOS Function and Sport Scores. *Phys Ther*. 2014;94(11):1614-1621. doi:10.2522/ptj.20140086
13. Weigl M, Wild H. European validation of The Comprehensive International Classification of Functioning, Disability and Health Core Set for Osteoarthritis from the perspective of patients with osteoarthritis of the knee or hip. *Disabil Rehabil*. 2018;40(26):3104-3112. doi:10.1080/09638288.2017.1377295
14. 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 Orthop*. 2009;80(1):55-61. doi:10.1080/17453670902805007
15. Harris K, Lim CR, Dawson J, Fitzpatrick R, Beard DJ, Price AJ. The Oxford knee score and its subscales do not exhibit a ceiling or a floor effect in knee arthroplasty patients: an analysis

of the National Health Service PROMs data set. *Knee Surgery, Sport Traumatol Arthrosc*. 2017;25(9):2736-2742. doi:10.1007/s00167-015-3788-0

16. Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in the OHS: An analysis of the NHS PROMs data set. *BMJ Open*. 2015;5(7). doi:10.1136/bmjopen-2015-007765
17. Behrend H, Giesinger K, Giesinger JM, Kuster MS. The "Forgotten Joint" as the Ultimate Goal in Joint Arthroplasty. Validation of a New Patient-Reported Outcome Measure. *J Arthroplasty*. 2012. doi:10.1016/j.arth.2011.06.035
18. Talbot S, Hooper G, Stokes A, Zordan R. Use of a New High-Activity Arthroplasty Score to Assess Function of Young Patients With Total Hip or Knee Arthroplasty. *J Arthroplasty*. 2010;25(2):268-273. doi:10.1016/j.arth.2008.09.019
19. Dawson J, Beard DJ, McKibbin H, Harris K, Jenkinson C, Price A. Development of a patient-reported outcome measure of activity and participation (the OKSAPQ) to supplement the Oxford knee score. *Bone Jt J*. 2014;96-B:332-338. doi:10.1302/0301-620X.96B3.32845
20. Bone A, March JJ, Dawson J, Beard DJ, Dawson J, Jenkinson C. Development of a patient-reported outcome measure of activity and participation (the OKSAPQ) to supplement the Oxford knee score Development of a patient-reported outcome measure of activity and participation (the OKS- APQ) to supplement the Oxford kn. *Bone Jt J*. 2016;96-B(March 2014):332-338. doi:10.1302/0301-620X.96B3.32845
21. Sahlqvist S, Song Y, Bull F, Adams E, Preston J, Ogilvie D. Effect of questionnaire length, personalisation and reminder type on response rate to a complex postal survey: Randomised controlled trial. *BMC Med Res Methodol*. 2011;11. doi:10.1186/1471-2288-11-62
22. Roorda LD, Crins MH, Terwee CB. Clinimetrics: Patient-Reported Outcomes Measurement Information System (PROMIS ®). *J Physiother*. 2019;65(2):110. doi:10.1016/j.jphys.2018.11.009
23. Terwee CB, Roorda LD, De Vet HCW, et al. Dutch-Flemish translation of 17 item banks from the Patient-Reported Outcomes Measurement Information System (PROMIS). *Qual Life Res*. 2014;23(6):1733-1741. doi:10.1007/s11136-013-0611-6
24. Crins MHP, van der Wees PJ, Klausch T, van Dulmen SA, Roorda LD, Terwee CB. Psychometric properties of the PROMIS Physical Function item bank in patients receiving physical therapy. *PLoS One*. 2018;13(2):1-14. doi:10.1371/journal.pone.0192187
25. Padilla JA, Rudy HL, Gabor JA, et al. Relationship Between the Patient-Reported Outcome Measurement Information System and Traditional Patient-Reported Outcomes for Osteoarthritis. *J Arthroplasty*. 2019;34(2):265-272. doi:10.1016/j.arth.2018.10.012
26. Stratford PW, Kennedy DM. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. *J Clin Epidemiol*. 2006;59(2):160-167. doi:10.1016/j.jclinepi.2005.07.012
27. Tolk JJ, Janssen RPA, Prinsen CAC, et al. The OARSI core set of performance-based measures for knee osteoarthritis is reliable but not valid and responsive. *Knee Surgery, Sport Traumatol Arthrosc*. 2019;27(9):2898-2909. doi:10.1007/s00167-017-4789-y
28. Tolk JJ, Janssen RPA, Prinsen CAC, Van Der Steen MC, Bierma-Zeinstra SMA, Reijman M. Measurement properties of the OARSI core set of performance-based measures for hip osteoarthritis: a prospective cohort study on reliability, construct validity and responsiveness in 90 hip osteoarthritis patients. *Acta Orthop*. 2018;90(1):15-20. doi:10.1080/17453674.2018.1539567
29. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. Cambridge: Cambridge University Press; 2011. doi:10.1017/CBO9780511996214
30. Mizner RL, Petterson SC, Clements KE, Zeni J a, Irrgang JJ, Snyder-Mackler L. Measuring functional improvement after total knee arthroplasty requires both performance-based

and patient-report assessments: a longitudinal analysis of outcomes. *J Arthroplasty*. 2011;26(5):728-737. doi:10.1016/j.arth.2010.06.004

31. Biggs PR, Whatling GM, Wilson C, Holt CA. Correlations between patient-perceived outcome and objectively-measured biomechanical change following Total Knee Replacement. *Gait Posture*. 2019;70(February):65-70. doi:10.1016/j.gaitpost.2019.02.028
32. Stevens-Lapsley JE, Schenkman ML, Dayton MR. Comparison of self-reported knee injury and osteoarthritis outcome score to performance measures in patients after total knee arthroplasty. *PM R*. 2011;3(6):541-549; quiz 549. doi:10.1016/j.pmrj.2011.03.002
33. Wiering B, de Boer D, Delnoij D. Asking what matters: The relevance and use of patient-reported outcome measures that were developed without patient involvement. *Heal Expect*. 2017;20(6):1330-1341. doi:10.1111/hex.12573
34. Alokozai A, Jayakumar P, Bozic KJ. Value-based Healthcare: Improving Outcomes through Patient Activation and Risk Factor Modification. *Clin Orthop Relat Res*. 2019;477(11):2418-2420. doi:10.1097/CORR.0000000000000985
35. Kaltoft MK, Dowie J. PROMs need PRIMs: Standardised Outcome Measures Lack the Preference-Sensitivity Needed in Person-Centred Care. *Stud Health Technol Inform*. 2019;262:118-121. doi:10.3233/SHTI190031
36. Hamilton DF, Lane J V, Gaston P, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open*. 2013;3(4):1-8. doi:10.1136/bmjopen-2012-002525
37. Noble PC, Conditt M a, Cook KF, Mathis KB. The John Insall Award: Patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452(452):35-43. doi:10.1097/01.blo.0000238825.63648.1e
38. 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):148-152. doi:10.1302/0301-620x.95b11.32767
39. Achaval S De, Kallen MA, Amick B, et al. Patients ' expectations about total knee arthroplasty outcomes. *Heal Expect*. 2016;19(2):299-308. doi:10.1111/hex.12350
40. Edwards PK, Mears SC, Lowry Barnes C. Preoperative Education for Hip and Knee Replacement: Never Stop Learning. *Curr Rev Musculoskelet Med*. 2017;10(3):356-364. doi:10.1007/s12178-017-9417-4
41. Mancuso C a, Sculco TP, Wickiewicz TL, et al. Patients' expectations of knee surgery. *J Bone Joint Surg Am*. 2001;83-A(7):1005-1012.
42. Mancuso CA, Salvati EA, Johanson NA, Peterson MGE, Charlson ME. Patients' expectations and satisfaction with total hip arthroplasty. *J Arthroplasty*. 1997;12(4):387-396. doi:10.1016/S0883-5403(97)90194-7
43. Smith EJ, Soon V-L, Boyd A, McAllister J, Deakin A, Sarungi M. What do Scottish Patients Expect of Their Total Knee Arthroplasty? *J Arthroplasty*. 2016;31(4):786-792. doi:10.1016/j.arth.2015.10.021
44. van den Akker-Scheek I, van Raay JJ a M, Reininga IHF, Bulstra SK, Zijlstra W, Stevens M. Reliability and concurrent validity of the Dutch hip and knee replacement expectations surveys. *BMC Musculoskelet Disord*. 2010;11(1):242. doi:10.1186/1471-2474-11-242
45. Macario A, Schilling P, Rubio R, Bhalla A, Goodman S. What questions do patients undergoing lower extremity joint replacement surgery have? *BMC Health Serv Res*. 2003;3(1):11. doi:10.1186/1472-6963-3-11
46. Rudan JF, Harrison MM, Grant HJ. Determining Patient Concerns Before Joint Arthroplasty. *J Arthroplasty*. 2009;24(7):1115-1119. doi:10.1016/j.arth.2008.08.007
47. Ekman I, Swedberg K, Taft C, et al. Person-centered care - Ready for prime time. *Eur J Cardiovasc Nurs*. 2011;10(4):248-251. doi:10.1016/j.ejcnurse.2011.06.008

48. Lingard EA, Sledge CB, Learmonth ID. Patient expectations regarding total knee arthroplasty: differences among the United States, United kingdom, and Australia. *J Bone Jt Surg Am*. 2006. doi:10.2106/jbjs.e.00147
49. Singh JA, Lewallen DG. Medical and psychological comorbidity predicts poor pain outcomes after total knee arthroplasty. *Rheumatology (Oxford)*. 2013;52:916-923. doi:10.1093/rheumatology/kes402
50. Singh JA, O'Byrne M, Harmsen S, Lewallen D. Predictors of moderate-severe functional limitation after primary Total Knee Arthroplasty (TKA): 4701 TKAs at 2-years and 2935 TKAs at 5-years. *Osteoarthr Cartil*. 2010;18(4):515-521. doi:10.1016/j.joca.2009.12.001
51. Edwards RR, Haythornthwaite JA, Smith MT, Klick B, Katz JN. Catastrophizing and depressive symptoms as prospective predictors of outcomes following total knee replacement. *Pain Res Manag*. 2009;14(4):307-311.
52. Hirschmann MT, Testa E, Amsler F, Friederich NF. The unhappy total knee arthroplasty (TKA) patient: Higher WOMAC and lower KSS in depressed patients prior and after TKA. *Knee Surgery, Sport Traumatol Arthrosc*. 2013;21(10):2405-2411. doi:10.1007/s00167-013-2409-z
53. Strine T, Chapman D, Kobau R, Balluz L, Mokdad A. Depression, Anxiety, and Physical Impairments and Quality of Life in the U.S. Noninstitutionalized Population. *Psychiatr Serv*. 2004;55(12):1408-1413.
54. Bourne RB, Chesworth B, Davis A, Mahomed N, Charron K. Comparing patient outcomes after THA and TKA: Is there a difference? In: *Clinical Orthopaedics and Related Research*. Vol 468. Springer New York; 2010:542-546. doi:10.1007/s11999-009-1046-9
55. Lau RL, Gandhi R, Mahomed S, Mahomed N. Patient Satisfaction after Total Knee and Hip Arthroplasty. *Clin Geriatr Med*. 2012;28(3):349-365. doi:10.1016/j.cger.2012.05.001
56. Tilbury C, Haanstra TM, Leichtenberg CS, et al. Unfulfilled Expectations After Total Hip and Knee Arthroplasty Surgery: There Is a Need for Better Preoperative Patient Information and Education. *J Arthroplasty*. 2016;31(10):2139-2145. doi:10.1016/j.arth.2016.02.061
57. Mancuso C, Graziano S, Briskie L, et al. Randomized trials to modify patients' preoperative expectations of hip and knee arthroplasties. *Clin Orthop Relat Res*. 2008;466(2):424-431. doi:10.1007/s11999-007-0052-z
58. Anakwe RE, Jenkins PJ, Moran M. Predicting dissatisfaction after total hip arthroplasty: a study of 850 patients. *J Arthroplasty*. 2011;26(2):209-213. doi:10.1016/j.arth.2010.03.013
59. Jain D, Nguyen L-CL, Bendich I, et al. Higher patient expectations predict higher patient-reported outcomes, but not satisfaction, in total knee arthroplasty patients: a prospective multicenter study. *J Arthroplasty*. 2017;32(9S):S166-S170. doi:10.1016/j.arth.2017.01.008
60. 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. doi:10.1016/j.arth.2017.07.021
61. Culliton SE, Bryant DM, MacDonald SJ, Hibbert KM, Chesworth BM. Effect of an e-Learning Tool on Expectations and Satisfaction Following Total Knee Arthroplasty: A Randomized Controlled Trial. *J Arthroplasty*. 2018;33(7):2153-2158. doi:10.1016/j.arth.2018.02.040
62. Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H. The role of expectancies in the placebo effect and their use in the delivery of health care: A systematic review. *Health Technol Assess (Rockv)*. 1999;3(3):1-96. doi:10.3310/hta3030
63. Ebrahim S, Malachowski C, Kamal el Din M, et al. Measures of Patients' Expectations About Recovery: A Systematic Review. *J Occup Rehabil*. 2014;25(1):240-255. doi:10.1007/s10926-014-9535-4
64. Flood AB, Lorence DP, Ding J, McPherson K, Black NA. The role of expectations in patients' reports of post-operative outcomes and improvement following therapy. *Med Care*. 1993;31(11):1043-1056. doi:10.1097/00005650-199311000-00006

65. Colloca L, Miller FG. Harnessing the placebo effect: the need for translational research. *Philos Trans R Soc Lond B Biol Sci*. 2011;366(1572):1922-1930. doi:10.1098/rstb.2010.0399
66. Morrell DC, Evans ME, Morris RW, Roland MO. The "five minute" consultation: Effect of time constraint on clinical content and patient satisfaction. *Br Med J (Clin Res Ed)*. 1986;292(6524):870. doi:10.1136/bmj.292.6524.870
67. Musellim B, Borekci S, Uzan G, et al. What should be the appropriate minimal duration for patient examination and evaluation in pulmonary outpatient clinics? *Ann Thorac Med*. 2017;12(3):177-182. doi:10.4103/atm.ATM_396_16
68. Chen LM, Farwell WR, Jha AK. Primary care visit duration and quality: Does good care take longer? *Arch Intern Med*. 2009;169(20):1866-1872. doi:10.1001/archinternmed.2009.341
69. Kortlever JTP, Ottenhoff JSE, Vagner GA, Ring D, Reichel LM. Visit Duration Does Not Correlate with Perceived Physician Empathy. *J Bone Jt Surg - Am Vol*. 2019;101(4):296-301. doi:10.2106/JBJS.18.00372
70. Bayne H, Neukrug E, Hays D, Britton B. A comprehensive model for optimizing empathy in person-centered care. *Patient Educ Couns*. 2013;93(2):209-215. doi:10.1016/j.pec.2013.05.016
71. McGuire LC. Remembering what the doctor said: Organization and adults' memory for medical information. *Exp Aging Res*. 1996;22(4):403-428. doi:10.1080/03610739608254020
72. Kessels RPC. Patients' memory for medical information. *J R Soc Med*. 2003;96(5):219-222. doi:10.1258/jrsm.96.5.219
73. Timmers T, Janssen L, Pronk Y, et al. Assessing the efficacy of an educational smartphone or tablet app with subdivided and interactive content to increase patients' medical knowledge: Randomized controlled trial. *JMIR mHealth uHealth*. 2018;6(12). doi:10.2196/10742
74. Sandberg EH, Sharma R, Sandberg WS. Deficits in retention for verbally presented medical information. *Anesthesiology*. 2012;117(4):772-779. doi:10.1097/ALN.0b013e31826a4b02
75. Dowsey MM, Spelman T, Choong PFM. Development of a Prognostic Nomogram for Predicting the Probability of Nonresponse to Total Knee Arthroplasty 1 Year After Surgery. *J Arthroplasty*. 2015;31:1654-1660. doi:10.1016/j.arth.2016.02.003
76. Chen JH, Asch SM. Machine learning and prediction in medicine-beyond the peak of inflated expectations. *N Engl J Med*. 2017;376(26):2507-2509. doi:10.1056/NEJMp1702071
77. Obermeyer Z, Emanuel E. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med*. 2016;375(13):1216-1219. doi:10.1056/NEJMp1609300
78. Taylor RA, Pare JR, Venkatesh AK, et al. Prediction of In-hospital Mortality in Emergency Department Patients with Sepsis: A Local Big Data-Driven, Machine Learning Approach. *Acad Emerg Med*. 2016;23(3):269-278. doi:10.1111/acem.12876
79. Wellner B, Grand J, Canzone E, et al. Predicting Unplanned Transfers to the Intensive Care Unit: A Machine Learning Approach Leveraging Diverse Clinical Elements. *JMIR Med Informatics*. 2017;5(4):e45. doi:10.2196/medinform.8680
80. Cabitza F, Locoro A, Banfi G. Machine Learning in Orthopedics: A Literature Review. *Front Bioeng Biotechnol*. 2018;6(June). doi:10.3389/fbioe.2018.00075

Summary

This thesis reports on outcome assessment and expectations of treatment result in total knee arthroplasty (TKA) and total hip arthroplasty (THA) patients, and whether expectation management can influence satisfaction after knee arthroplasty.

Part 01 | Outcome assessment after arthroplasty

The first part of this thesis focuses on measurement of physical function after joint arthroplasty. This part consists of 2 studies on the measurement properties of the recommended set of performance-based measures (PBMs) to assess physical function in knee and hip osteoarthritis patients. The core set consists of the 30-second chair-stand test (30s CST), 4x10 meter fast-paced walk test (40m FPWT) and a stair climb test. These tests aim to asses physical function in the domains sit-to-stand movement, walking short distances and stair negotiation.

In **Chapter 2** the reliability, validity and responsiveness of the OARSI core set of PBMs was evaluated in a prospective cohort study of 85 knee osteoarthritis patients indicated for total knee arthroplasty (TKA). Construct validity and responsiveness were assessed by testing of predefined hypotheses, and test-retest measurements for reliability analysis. Patient reported outcome measures (PROMs) on pain, physical function and quality of life, and knee extensor strength were used as comparator instruments. Measurements were obtained at baseline and 12 months after TKA. Appropriate test-retest reliability was found for all three tests (Intraclass Correlation Coefficients (ICC) between 0.90 and 0.93). Adequate construct validity could not be confirmed for the three tests, as none of the tests reached 75% hypothesis confirmation. In the responsiveness analysis only the 40m FPWT was found to be responsive with 75% of predefined hypothesis confirmed. Based on these results it was concluded that the PBMs of the OARSI core set had good reliability, but poor construct validity and responsiveness. Therefore, these findings do not justify their use for clinical practice in patients with knee osteoarthritis.

In **Chapter 3** a similar approach was used to assess the reliability, validity and responsiveness of the OARSI core set of PBMs in a prospective cohort of 90 patients with hip osteoarthritis scheduled for total hip arthroplasty (THA). Test-retest measures were used for the reliability analysis and the hypothesis testing method for construct validity and responsiveness analysis. Test-retest reliability was appropriate (ICCs between 0.86 and 0.96), but in the construct validity and responsiveness analysis none of the performance-based measures reached 75% hypothesis confirmation. Therefore, it was concluded that the OARSI core set of PBMs are not suitable for clinical practice in patient with hip osteoarthritis either.

Part 02 | Expectations of treatment result

The second part of this thesis addresses what can be considered realistic expectations for treatment result after arthroplasty and analysed determinants of patients' expectations of treatment result.

In **chapter 4** we reported on what the members of the Dutch Knee Society assume are realistic expectations for recovery one year after TKA. All members were invited to fill out a web-based questionnaire, the response rate was 56%. Expectations were measured using the Dutch version of the Hospital for Special Surgery (HSS) knee replacement expectations survey. The overall HSS knee replacement expectation score was 66.0 (Standard Deviation [SD] 14.0) on a 0-100 scale. Most improvement was predicted on the domains pain, function, activities and psychological wellbeing. Nevertheless, complete return to normal should not be expected according to the respondent; especially in demanding physical activities residual symptoms are likely to occur.

In **chapter 5** a systematic review is presented, summarizing the current available evidence about when patients might resume driving after elective, primary THA or TKA. In February 2016 a systematic literature search was conducted, searching for clinical studies reporting on 'THA', 'TKA', 'car driving', 'reaction time' and 'brake response time'. Two researchers independently screened the titles and abstracts for eligibility and assessed the risk of bias. Both fixed and random effects were used to pool data and calculate differences between pre- and post-operative total brake response time (TBRT). A total of 19 studies were included, of these studies one was at high risk for bias, six studies at moderate risk and 12 studies at low risk. Meta-analysis showed that there was no difference in TBRT with the pre-operative values as from two weeks after right sided THA (mean decrease of 2.6 seconds [-3.2 to 8.3]), and four weeks after a right-sided TKA (mean decrease 1.8 seconds [-0.6 to 4.3]). These results can be used as general guidelines in advising patients when to resume driving after TKA or THA. However, it must be noted that the advice should be individualised.

The study presented in **chapter 6** analysed to what extent pre-operative outcome expectations of TKA patients are affected by psychological factors, demographic factors, pain, physical function and general health status. A cross-sectional analysis of 204 patients scheduled for primary TKA for knee osteoarthritis, was performed. The HSS knee replacement expectations survey was used to measure outcome expectations. Other factors measured were age, sex, body mass index and PROMs on pain, physical function, quality of life, anxiety, depression, catastrophizing, optimism and pessimism. Associations between these variables and pre-operative outcome expectations were assessed using multiple linear regression analyses. Female sex, higher age, higher depression score and duration of complaints > 50 months showed to be significant predictors of lower expectations for the treatment outcome after TKA. Baseline pain and function scores showed not to be related to the level of pre-operative expectations. These results are useful for improving and individualising expectation management interventions, by identifying patients at risk for having either too high or too low expectations.

Part 03 | Expectation management in clinical practice

In the third part of this thesis we aimed to translate the potential of improved expectation management to clinically applicable interventions.

The study protocol for a randomised controlled trial (RCT) that assesses the effect of an additional expectation management module for TKA patients is presented in **chapter 7**. The primary objective of the presented study was to examine whether an educational module on long-term recovery after TKA would improve patient satisfaction compared to usual pre-operative education. Patients with symptomatic and radiographic knee osteoarthritis scheduled for a primary, unilateral TKA were randomized to usual pre-operative education (control group) or usual education plus an additional module on realistic expectations for long-term recovery (intervention group). Observers were blinded to group allocation and patients naïve to study objective and difference between study groups. Outcome expectations were measured using the HSS Knee Replacement Expectations Survey at baseline (before the intervention), pre-operative (after the intervention) and fulfilment of expectations at 12-month follow-up. The primary outcome was being very satisfied (numerical rating scale for satisfaction ≥8) with the treatment result at 12-month follow-up. Analyses were performed intention-to-treat according to group allocation and per-protocol for patients that attended the education sessions.

Chapter 8 presents the results of this RCT. Between July 2016 and April 2018, 459 consecutive patients with knee osteoarthritis indicated for primary, unilateral TKA were assessed for eligibility. 204 patients were randomized to either usual pre-operative education or usual education plus an additional module on realistic expectations for long-term recovery. 187 patients (91.7%) were available for analysis at follow-up. In the intention-to-treat analysis 69.9% patients were very satisfied with the treatment result in the intervention group, and 58.5% patients in the control group, between group difference 11.4% (95%CI -2.4 – 25.2%). A per-protocol analysis for patients that attended the education session (94.1%) showed 74.4% very satisfied patients in the intervention group and 56.9% in the control group (mean difference 17.4% [95%CI 3.3-31.6%]). After pre-operative education the expectation scores in the intervention group were significantly lower (mean difference -6.9 [95%CI -10.2 – -3.6]) and unchanged in the control group (mean difference 0.5 [95%CI -2.9 – 3.9). Overall fulfilment of expectations at 12-months follow-up was 70% (SD 28.8) in the intervention group and 58.6% (SD 33.0) in the control group; mean difference 11.4% (95%CI 2.3-20.5). We concluded that improved pre-operative patient education modified patient expectations, this resulted in higher postoperative expectation fulfilment and higher postoperative satisfaction.

The development and validation of prediction models for residual symptoms on 10 specific outcome parameters at 12-month follow-up for patients undergoing primary TKA for knee osteoarthritis is presented in **chapter 9**. For development of the models, routinely gathered demographic and PROMs data was extracted from the Dutch Arthroplasty Register (Landelijke Registratie Orthopedische Implantaten [LROI]) on 7071 patients with a TKA. Multiple logistic regression analyses were performed to construct predictive algorithms for satisfaction, treatment success, and residual symptoms concerning pain in rest and during activity, sit-to-stand movement, stair negotiation, walking, performance of activities of

daily living, kneeling and squatting. Residual complaints on kneeling (♀ 72% / ♂ 59%) and squatting (♀ 71% / ♂ 56%) were reported most frequently, and least residual complaints were scored for walking (♀ 16% / ♂ 12%) and pain in rest (♀ 18% / ♂ 14%). The predictive algorithms for residual symptoms concerning sit-to-stand movement, stair negotiation, walking, activities of daily living and treatment success showed acceptable discriminative values (AUC 0.68 – 0.74). The prediction models for residual complaints regarding kneeling, squatting, pain and satisfaction showed the least favourable results (AUC 0.58 – 0.64). The calibration curves showed adequate calibration for most of the models. We concluded that demographic and PROMs data collected for the LROI registry are usable for outcome prediction, this prediction is useful for individual expectation management in patients planned for TKA.

Finally, in **Chapter 10** a general discussion is presented on the main findings, future research perspectives and implications for clinical practice of the studies described in this thesis.

Samenvatting

Dit proefschrift rapporteert over uitkomstmeting en verwachtingen van het behandelresultaat bij patiënten die een totale knieprothese (TKP) of totale heupprothese (THP) operatie ondergaan en of verwachtingsmanagement de tevredenheid na een TKP kan beïnvloeden.

Deel 01 | Uitkomstmeting na heup en knie prothesiologie

Het eerste deel van dit proefschrift richt zich op het meten van fysiek functioneren na een gewrichtsprothese. Dit deel bestaat uit 2 studies naar de meeteigenschappen van de door de Osteoarthritis Reasearch Society International (OARSI) aanbevolen set functietesten bij patiënten met knie- of heupartrose. De 'core set' bestaat uit de '30-second chair-stand test' (30 s CST), '4x10 meter fast-paced walk test' (40 m FPWT) en een test die traplopen beoordeelt. Deze testen beogen fysiek functioneren te meten in de domeinen opstaan uit een stoel, het lopen van korte afstanden en traplopen.

Hoofdstuk 2 beschrijft een evaluatie van de betrouwbaarheid, validiteit en responsiviteit van de functietesten uit de OARSI core set voor knieartrose, in een prospectieve cohortstudie bij 85 patiënten met knieartrose, geïndiceerd voor een TKP. Constructvaliditeit en responsiviteit werden beoordeeld door het toetsen van vooraf gedefinieerde hypothesen en een analyse van de betrouwbaarheid werd verricht met test-hertest metingen. Quadricepskracht en vragenlijsten (patient reported outcome measures [PROMs]) ten aanzien van pijn, fysiek functioneren en kwaliteit van leven werden als vergelijkende instrumenten gebruikt. Metingen werden verricht voor de behandeling en 12 maanden na de TKP-operatie. Er werd een goede test-hertest betrouwbaarheid gevonden voor alle drie de testen (intraclass correlatiecoëfficiënten [ICC] tussen 0,90 en 0,93). De constructvaliditeit van de drie testen kon niet worden bevestigd, aangezien bij geen van de testen meer dan 75% van de hypotheses werd bevestigd. In de responsiviteitsanalyse liet alleen de 40m FPWT goede resultaten zien; 75% van de vooraf gedefinieerde hypothesen werd bevestigd. Op basis van deze resultaten werd geconcludeerd dat de functietesten van de OARSI core set een goede betrouwbaarheid hadden, maar dat de constructvaliditeit en responsiviteit niet kon worden bevestigd. Gezien deze bevindingen is gebruik van deze testen voor de klinische praktijk bij patiënten met knieartrose niet gerechtvaardigd.

In **hoofdstuk 3** werd een vergelijkbare aanpak gebruikt om de betrouwbaarheid, validiteit en responsiviteit van de functietesten van de OARSI core set voor heupartrose te beoordelen in een prospectieve cohortstudie bij 90 patiënten met heupartrose die waren gepland voor een THP. Test-hertest metingen werden gebruikt voor de betrouwbaarheidsanalyse en voor de constructvaliditeits- en responsiviteitsanalyse werd de hypothese toetsingsmethode gebruikt. De studie toonde dat de test-hertest betrouwbaarheid goed was (ICC's tussen 0,86 en 0,96). In de constructvaliditeit- en responsiviteitsanalyse werd bij geen enkele van de functietesten 75% of meer van de hypotheses bevestigd. Daarom werd geconcludeerd dat de functietesten van de OARSI core set niet geschikt zijn voor toepassing in de klinische praktijk bij patiënten met heupartrose.

Deel 02 | Verwachtingen van het resultaat van de behandeling

Het tweede deel van dit proefschrift gaat in op wat realistische verwachtingen zijn voor het behandelresultaat na een gewricht vervangende prothese en welke factoren van invloed zijn op de hoogte van deze verwachtingen.

In **hoofdstuk 4** rapporteerden we wat de leden van de NOV werkrgoep knie (Dutch Knee Society) realistische verwachtingen vinden voor het te verwachten herstel één jaar na een TKP operatie. Alle leden werden uitgenodigd om een digitale vragenlijst in te vullen, het responspercentage was 56%. De verwachtingen werden gemeten met behulp van de Nederlandse versie van de Hospital for Special Surgery (HSS) knee replacement expectations survey. De respondenten rapporteerden en gemiddelde HSS knee replacement expectations totaalscore van 66,0 (Standard Deviatie [SD] 14,0) op een schaal van 0-100. De meeste verbetering werd voorspeld voor de domeinen pijn, functie, activiteiten en psychologisch welzijn. Desalniettemin mag volgens de respondenten geen volledige terugkeer naar normaal worden verwacht; vooral bij veeleisende fysieke activiteiten zijn restklachten te verwachten.

Een systematische review naar wanneer patiënten weer kunnen autorijden na een electieve, primaire THP of TKP operatie wordt gepresenteerd in **hoofdstuk 5**. Hierin werd het in de literatuur beschikbare bewijs samengevat. In februari 2016 werd een systematisch literatuuronderzoek verricht, waarbij werd gezocht naar klinische studies die rapporteerden over 'THA', 'TKA', 'car driving', 'reaction time' en 'brake response time'. Twee onderzoekers hebben onafhankelijk van elkaar de titels en samenvattingen gescreend op geschiktheid en het risico op bias beoordeeld. Zowel fixed- als random effect models werden gebruikt om gegevens te bundelen en verschillen te berekenen tussen pre- en postoperatieve totale rem-reactietijd (total brake response time [TBRT]). In totaal werden 19 studies geïncludeerd, van deze studies was er één met een hoog risico op bias, zes studies met een matig risico en 12 studies met een laag risico. Meta-analyse toonde aan dat er geen verschil was in TBRT met de preoperatieve waarden vanaf twee weken na rechtszijdige THP (gemiddelde 2,6 seconden [-3,2 tot 8,3] korter) en vier weken na een rechtszijdige TKP (gemiddelde 1,8 seconden [-0,6 tot 4,3] korter). Deze resultaten kunnen worden gebruikt als algemene richtlijnen voor het adviseren van patiënten wanneer ze na TKP en THP weer kunnen autorijden. Daarbij moet aangemerkt worden dat een advies altijd aangepast moet worden aan individuele omstandigheden.

In de studie gepresenteerd in **hoofdstuk 6** werd geanalyseerd in hoeverre preoperatieve uitkomstverwachtingen van patiënten die op de wachtlijst staan voor een TKP worden beïnvloed door psychologische factoren, demografische factoren, pijn, fysiek functioneren en de algemene gezondheidstoestand. Er werd een cross-sectionele analyse verricht bij 204 patiënten die op de wachtlijst stonden voor een primaire TKP vanwege knieartrose. De HSS knee replacement expectations survey werd gebruikt om verwachtingen te meten. Andere factoren die gemeten werden waren leeftijd, geslacht, body mass index en PROMs over pijn, fysiek functioneren, kwaliteit van leven, angst, depressie, catastroferen, optimisme en pessimisme. Associaties tussen deze variabelen en preoperatieve uitkomstverwachtingen werden beoordeeld met behulp van multipele lineaire regressieanalyses. Vrouwelijk geslacht, hogere leeftijd, hogere depressiescore en duur van klachten langer dan 50

maanden bleken significante voorspellers te zijn van lagere verwachtingen voor het behandelresultaat na een TKP. Pijn- en functiescores bleken niet gerelateerd te zijn aan het niveau van preoperatieve verwachtingen. Deze resultaten zijn bruikbaar bij het verbeteren en individualiseren van interventies op het gebied van verwachtingsmanagement. Hiermee kunnen namelijk patiënten die een grotere kans hebben op te hoge of te lage verwachtingen geïdentificeerd worden.

Deel 03 | Verwachtingsmanagement in de klinische praktijk

Het derde deel van dit proefschrift had als doel om het potentieel van verbeterd verwachtingsmanagement te vertalen naar klinisch toepasbare interventies.

In **hoofdstuk 7** wordt het studieprotocol gepresenteerd voor een gerandomiseerde gecontroleerde studie (Randomised Controlled Trial [RCT]) die het effect van een aanvullende verwachtingsmanagement module bij TKP-patiënten beoordeelt. Doel van de studie was te beoordelen of een aanvullende voorlichtingsmodule over lange termijn herstel na een TKP leidt tot een hogere postoperatieve patiënt tevredenheid in vergelijking met de gebruikelijke preoperatieve voorlichting. Patiënten met symptomatische en radiografische knieartrose gepland voor een primaire, unilaterale TKA werden gerandomiseerd naar ofwel de reguliere preoperatieve voorlichting (controlegroep) of de reguliere voorlichting plus een aanvullende module over realistische verwachtingen voor lange termijn herstel (interventiegroep). De onderzoekers waren geblindeerd voor groepsallocatie en patiënten waren naïef voor het doel van de studie en het verschil tussen de twee voorlichtingsmodules. Verwachtingen van het behandelresultaat werden gemeten met de HSS Knee Replacement Expectations Survey op baseline (vóór de interventie), preoperatief (na de interventie) en vervulling van verwachtingen werd gemeten bij 12 maanden follow-up. De primaire uitkomst was of een patiënt zeer tevreden (numerical rating scale [NRS] voor tevredenheid ≥8) was met het resultaat van de behandeling na 12 maanden follow-up. Analyses werden intention-to-treat uitgevoerd volgens gerandomiseerde groep en een per-protocol analyse werd uitgevoerd voor patiënten die daadwerkelijk de voorlichtingssessie bijwoonden.

Hoofdstuk 8 beschrijft de resultaten van deze RCT. Tussen juli 2016 en april 2018 werd bij 459 opeenvolgende patiënten met knieartrose, geïndiceerd voor primaire, unilaterale TKP, beoordeeld of ze in aanmerking kwamen voor deelname aan de studie. 204 patiënten werden gerandomiseerd naar ofwel reguliere preoperatieve voorlichting, ofwel reguliere voorlichting plus een aanvullende module over realistische verwachtingen voor lange termijn herstel. 187 patiënten (91,7%) waren bij follow-up beschikbaar voor analyse. De intention-to-treat-analyse toonde dat in de interventiegroep 69,9% van de patiënten zeer tevreden was met het resultaat van de behandeling, in de controlegroep was dit 58,5; gemiddeld verschil tussen de groepen 11,4% (95% BI -2,4 - 25,2%). Een per-protocol-analyse voor patiënten die daadwerkelijk de voorlichting bijwoonden (94,1%) toonde 74,4% zeer tevreden patiënten in de interventiegroep en 56,9% in de controlegroep (gemiddeld verschil 17,4% [95% BI 3,3-31,6%]). Na preoperatieve voorlichting waren de verwachtingsscores gemeten met de HSS Knee Replacement Expectations Survey in de interventiegroep significant lager (gemiddeld verschil -6,9 [95% BI -10,2 - -3,6]) en ongewijzigd in de

controlegroep (gemiddeld verschil 0,5 [95% BI -2,9 - 3,9) . Het percentage uitgekomen verwachtingen bij 12 maanden follow-up was 70% (SD 28.8) in de interventiegroep en 58.6% (SD 33.0) in de controlegroep; gemiddeld verschil 11,4% (95% BI 2,3-20,5). Deze resultaten leidden tot de conclusie dat verbeterde preoperatieve voorlichting de verwachtingen van de patiënt veranderde, hetgeen resulteerde in een hoger percentage uitgekomen postoperatieve verwachting en uiteindelijk een hogere postoperatieve tevredenheid bij de patiënten die de voorlichtingssessie hadden bezocht.

De ontwikkeling en validatie van voorspellingsmodellen voor restklachten op 10 specifieke uitkomstparameters 12 maanden na een primaire TKA wordt gepresenteerd in **hoofdstuk 9**. Voor de ontwikkeling van de modellen werden routinematig verzamelde demografische data en PROM-scores van 7071 TKP-patiënten geëxtraheerd uit de Landelijke Registratie Orthopedische Implantaten [LROI]. Multipele logistische regressieanalyses werden uitgevoerd om voorspellende algoritmen op te stellen voor tevredenheid, behandelsucces en resterende symptomen ten aanzien van pijn in rust en tijdens activiteit, opstaan uit een stoel, traplopen, lopen, algemene dagelijkse levensverrichtingen, knielen en hurken. Restklachten bij knielen (♀ 72% / ♂ 59%) en hurken (♀ 71% / ♂ 56%) werden het vaakst gerapporteerd en de minste restklachten werden gescoord voor wandelen (♀ 16% / ♂ 12%) en pijn in rust (♀ 18% / ♂ 14%). De voorspellende algoritmen voor resterende symptomen bij opstaan uit een stoel, traplopen, wandelen, algemene dagelijkse levensverrichtingen en behandelsucces toonden acceptabele discriminatie (AUC 0,68 - 0,74). De voorspellingsmodellen voor restklachten bij knielen, hurken, pijn en tevredenheid presteerden minder goed (AUC 0,58 - 0,64). De kalibratiecurven toonden voor de meeste modellen adequate kalibratie. Op basis van deze resultaten concludeerden wij dat demografische en PROMs-gegevens die routinematig worden verzameld de voor de LROI, bruikbaar zijn voor het voorspellen van de uitkomst na een TKP. Deze voorspelling is bruikbaar voor geïndividualiseerd verwachtingsmanagement bij patiënten gepland voor een TKP.

Ten slotte wordt in **hoofdstuk 10** een algemene discussie gepresenteerd over de belangrijkste bevindingen, toekomstige onderzoeksperspectieven en implicaties voor de klinische praktijk van de studies die in dit proefschrift zijn beschreven.

Dankwoord

De vele patiënten die bereid waren om deel te nemen aan de verschillende studies in dit proefschrift verdienen uiteraard het eerste woord van dank. Zonder klinisch onderzoek geen vooruitgang in patiëntenzorg, maar bovenal zonder patiënten geen klinisch onderzoek. Dankzij uw deelname kunnen we toekomstige patiënten nog beter behandelen.

Prof. dr. Bierma-Zeinstra, beste Sita. In vergelijking met de eerste opzet die we voor dit proefschrift maakten is in de eindversie vrijwel geen enkel hoofdstuk hetzelfde geworden. Mede dankzij jouw scherpe analyses en zicht op de grote lijnen is het een samenhangend geheel geworden. Ik vond het indrukwekkend te zien hoe je met een paar rake observaties de kern van elk stuk wist te raken, met duidelijke adviezen om de boodschap aan te scherpen en nog helderder te presenteren. Veel dank voor je begeleiding als promotor.

Dr. Reijman, beste Max. Zonder jouw intensieve betrokkenheid als copromotor was dit proefschrift niet tot stand gekomen. Dat geldt overigens ook voor het hoogtepunt van onze wetenschappelijk samenwerking *'De invloed van de hond versus de mens Max op wetenschappelijke output en werkplezier'*, dat vreemd genoeg dit boek toch niet heeft gehaald. Daarnaast heb ik methodologisch gezien veel van je geleerd, ook dat statistische analyses uiterst bruikbaar zijn buiten de wetenschap. Zo weet ik tegenwoordig dat het goed mogelijk is om niet als eerste boven te zijn, maar toch een veel betere wielrenner te zijn. Het blijkt namelijk dat de uitslag altijd gecorrigeerd moet worden voor jouw leeftijdsgerelateerde wattageverlies. Naast deze grappen heb ik jouw gedrevenheid en kritische ondersteuning ontzettend gewaardeerd. Wellicht komen er in de toekomst nog gezamenlijke (kinder-) orthopedische projecten.

Dr. Janssen, beste Rob. Jij hebt me vanaf de eerste wetenschappelijke stappen in Máxima MC begeleidt. Dit leidde ertoe dat je mijn copromotor werd, en gaandeweg mijn opleiding tot orthopedisch chirurg werd je ook mijn opleider. Ik heb dan ook niet alleen op het gebied van onderzoek veel van je opgestoken, maar ook klinisch belangrijke lessen geleerd. Ik vind het inspirerend om te zien hoe je, zonder concessies aan de kwaliteit van je werk, op zo veel vlakken actief kunt zijn. Hoewel ik jouw stelling dat leuk alleen voor thuis is niet volledig onderschrijf (maar ik behandel dan ook niet alleen knieën…), is gelukkig het grootste deel van dit proefschrift thuis geschreven. Dus daar kunnen we het zeker over eens zijn. Ook buiten het vak heb ik veel van je geleerd, met name *'la nourriture est la base de la vie'* zal ik nooit vergeten. Veel dank Rob, *'weg met de terreur van de middelmaat'*; het is mede dankzij jou een prachtig proefschrift geworden.

Geachte leden van de promotiecommissie: Prof.dr. L.J.W. van Rhijn, Prof. dr. B.W. Koes en Prof.dr. J.M.W. Hazes. Dank voor uw bereidheid en inzet om mijn wetenschappelijke werk op zijn merites te beoordelen. Ik kijk er naar uit om met u van gedachten te wisselen over de inhoud.

Verpleegkundig specialisten Christa van Doesburg en Hilke Cox. Jullie enthousiasme voor de dagelijkse zorg voor prothese patiënten en doorlopende motivatie om praktische verbeteringen hierin aan te brengen, heeft veel toegevoegd aan de ideeën voor de studies in dit proefschrift. Jullie bereidheid om bij te dragen aan de extra logistieke vereisten waren onmisbaar om deze ideeën ook wetenschappelijk te onderzoeken. Hartelijk dank voor deze samenwerking.

Dames van de poli orthopedie in Máxima MC: Corianne, Monique, Regina, Jolanda, Karin, Mariette, Judith, Miranda, Gina, Marion, Karin, Bea, Carin, Anne, Monique, Karin, Lia, Els en dames van de kinderorthopedie Karlijn, Dorien, Wendy, Anneke, Miriam, Yvon, Monique, Elly, Evi en Zoë. Jullie zagen (en zien) me regelmatig aankomen met een ideetje om iets te onderzoeken. Hoewel bijna alle studies in het reguliere zorgproces ingepast konden worden, was er wel steeds weer aandacht nodig om te organiseren dat de juiste patiënt de juiste informatie kreeg en ook nog op de goede plek aankwam. Veel dank voor jullie geduld en ondersteuning hierbij.

Mijn co-auteurs Marieke van der Steen, Erwin Waarsing, Tsjitske Haanstra, Charlotte van der Velden, Sanna Prinsen, Daisy Latijnhouwers en Liza van Steenbergen. Bedankt voor jullie onmisbare hulp bij het schrijven van de manuscripten voor dit proefschrift.

Stafleden Orthopedie Groot Eindhoven. Afdeling orthopedie en traumatologie Máxima MC: Jan van Mourik, Henk Koot, Willem den Boer, Rob Janssen, Florens van Douveren, Anouk Giesberts, Hans Hendriks, Marijn van de Besselaar, Inge Bonneux, Coen Jaspers, Arnold Besselaar en afdeling orthopedie Catharina ziekenhuis Eindhoven: Robin van Kempen, Niek Schepel, Janneke Bos, Remco van Wensen, Paul de Baat, Thijn Fuchs en in memoriam Kees Oosterbos. Jullie maakten mijn opleiding tot orthopedisch chirurg zoals hij was. Onderzoek dat dicht bij de klinische praktijk staat geeft veel ruimte voor discussie over het klinische perspectief en praktische inzichten. Jullie ervaring, ideeën en enthousiasme voor optimale patiëntenzorg staat aan de basis van dit proefschrift. Veel dank voor de intensieve begeleiding en ruimte die ik de afgelopen jaren hebben gekregen om me zowel klinisch als wetenschappelijk te kunnen ontwikkelen.

Een extra woord van dank voor dr. van Mourik, beste Jan. Dat de lat hoog lag bij mijn entree in het Máxima was wel duidelijk toen binnen korte tijd een krantenartikel getiteld ‘Tolk is de redding van de patiënt’ op het prikbord hing. Je zag erop toe dat ik er mijn best voor deed dat dat op zijn minst een beetje zo was, en gelukkig leidde dit ertoe dat je me aannam voor de opleiding tot orthopedisch chirurg. Het ongeëvenaarde opleidingklimaat dat jij hebt bewerkstelligd heeft er ongetwijfeld aan bijgedragen dat dit proefschrift (vrijwel) helemaal tijdens mij opleiding tot stand is kunnen komen, en ik ook een gedegen opleiding tot orthopedisch chirurg heb gekregen. Ik zie je rust, integriteit en wijsheid als voorbeeld. Dank dat ik veel van je heb mogen leren.

Beste Arnold en Florens, zoals beloofd zou ik expliciet benoemen dat jullie niets aan dit proefschrift hebben bijgedragen. Dat is maar deels waar. Knie en heupprothesen zijn dan wel niet jullie dagelijks werk, maar de concepten ten aanzien van verwachtingen en verwachtingsmanagement passen jullie bij kinderen dagelijks toe. En ondanks dat het misschien niet zo opvalt, let ik vaak goed op wat jullie zeggen en doen. Dank voor de mogelijkheid om me wat minder op protheses en meer op de kinderorthopedie te kunnen richten, wat hebben we toch een prachtig vak.

Beste collega AIOS orthopedie van ROGO Zuid, dank voor de mooie opleidingstijd die we samen hebben gehad. In het bijzonder Thijn Fuchs en Frank Jonkers, onze wegen kwamen samen in het Máxima en bleven dat tot ver daarbuiten. Hoewel we de liefde voor

de wetenschap wisselend delen, komt die voor het vak zeker overeen. We weten nog niet allemaal wat de toekomst ons brengt, maar dat we elkaar blijven vinden staat vast.

Paranimfen Sven Akkerman en Floris Jansen. Beste Sven, omdat ik meestal mijn tijd neem voor dingen heb ik bij jou kunnen afkijken hoe het hoort. Ik hoop dat ik het zonder roestige tapijtscharen heb kunnen evenaren. In ieder geval heb ik onthouden dat niet aflatende verwondering vaak tot de beste inzichten leidt. Floris, jouw plek als paranimf berust niet op statistische vakkennis of diepe methodologische inzichten die tot steun kunnen zijn tijdens de verdediging. Die heb je ongetwijfeld in huis, maar onder een sluier van dadaïstisch absurdisme heb ik daar nog niet alles van kunnen ontdekken. Wellicht dat een combinatie van jullie benaderingen de beste strategie is; *if you can't convince them, confuse them*. Dank dat jullie ook deze dag achter mij staan. Jullie zijn er bij de belangrijke en mooie dingen, samen met Rianne, Rhys, Hedda, Roelien, Julius en Ferdinand maken jullie het compleet.

Mannen mijn vrienden. Zo zie je maar waar een gedegen basis van mijnenveger, Ferrari dekbedden en gepaste sociale afstand toe kan leiden. Tijdens onze tijd in Maastricht ging veel zoals verwacht en nog veel meer niet, en dat was misschien wel het mooiste deel. Sindsdien, met en zonder fiets, blijft dat nog steeds precies hetzelfde. Ik kan overigens niet veel in dit proefschrift bedenken waar ik jullie direct voor zou moeten bedanken, maar gelukkig is er veel meer in het leven en daar spelen jullie zeker een belangrijke rol in.

Wim en Ineke. Naast oneindige dank voor jullie dochter, waardeer ik jullie oprechte interesse in de voortgang en resultaten van mijn onderzoek erg. Uiteraard minder leuk dan aandacht van en voor de dames, maar daar valt dan ook niet mee te concurreren. Het is fantastisch dat jullie er wekelijks voor ze zijn.

Lieve Lieselotte. Ik kijk er altijd naar uit om samen met jou, Erik en de mannen het leven te vieren. Onvergetelijke avonden. Ik vind het mooi om nu wetenschappelijk in jouw voetsporen te treden.

Ankie en Jan Maarten. Als het om verwachtingen gaat is het terecht om met zo'n goede basis veel van mij te verwachten. Ik weet zeker dat het jullie aanspreekt dat dit proefschrift meer over praten dan over snijden gaat, en zo in alle opzichten goed aansluit bij wat jullie zelf deden voordat de grote vakantie begon. Oneindig veel dank voor jullie warmte, liefde en onvoorwaardelijk vertrouwen in de goede afloop. Zonder jullie was het niet zoals het is, en het is goed.

Lieve Julia. *A million miles beyond what science understands*. Waar veel in dit proefschrift draait om wat de patiënt belangrijk vindt, is er in het echte leven natuurlijk maar 1 ding echt belangrijk; jij en de meisjes. Zonder jou was er niets van dit alles en zonder Olivia en Valerie deed het er ook niet toe. Jullie maken mij elke dag de gelukkigste man en vader die er bestaat.

Curriculum vitae

Jaap Tolk was born on the 13th of Januari 1983 in Venlo, the Netherlands. In 2001, he graduated at Collegium Marianum (VWO) and went on to study medicine at Maastricht University in the same year. During his medical training he gained a genuine interest in musculoskeletal pathology, and saw this interest acknowledged during an elective clinical rotation at the department of orthopedic surgery at VieCuri Medical Center in Venlo under supervision of dr. J.W. Morrenhof. He followed his last year clinical and research internships at Máxima MC in Eindhoven and Veldhoven, after which he had the opportunity to start as a resident in the same hospital.

His orthopedic training in ROGO Zuid started in 2013 at the department of general surgery in Catharina Hospital Eindhoven (dr. G.A.P. Nieuwenhuijzen). This was followed by orthopedic residencies at Viecuri MC (dr. J.W. Morrenhof and dr. F.O. Lambers Heerspink), Maastricht University Medical Center+ (prof.dr. L.W. van Rhijn) and completed in 2019 in Máxima MC (dr. J.B.A van Mourik and dr. R.P.A. Janssen). The research projects included in this thesis were initiated and concluded in Máxima MC within his orthopedic surgery training period.

In the course of his residency, he became particularly interested in pediatric orthopedic surgery. This resulted in differentiation in his senior residency year and subsequently his currently held position as fellow pediatric orthopedic surgery in Máxima MC (supervision F.Q.M.P. van Douveren and A.T. Besselaar). To further specialize in this field and gain international experience he will start as fellow pediatric orthopedic surgery in the Royal National Orthopedic Hospital in Stanmore, United Kingdom, under the supervision of prof. dr. D.M. Eastwood in July 2020.

Jaap is married to Julia Spierings and they have two daughters: Olivia and Valerie.

PhD portfolio

Summary of PhD training and teaching
PhD student: Drs. J.J. Tolk
PhD Period: 2014-2020
Erasmus University Medical Centre Rotterdam, Department of Orthopaedic Surgery
Promotor: Prof. dr. S.M.A Bierma-Zeinstra
Co-promotores: Dr. M. Reijman, Dr. R.P.A. Janssen

01. PhD training

General courses	year	ECTS
Good Clinical Practice	2011	1.0
Basic clinical teaching, Maastricht University Medical Centre	2014	0.5
Good Clinical Practice, certification update	2016	0.5
Clinical Epidemiology, Utrecht University	2017	1.0
Design and Interpretation of Clinical Trials, Johns Hopkins University	2018	1.0

Podium presentations	year	ECTS
Performance-based measures for knee osteoarthritis, evaluation of the OARSI core set. NOV autumn meeting, Veldhoven	2017	1.0
What determines patient expectations of total knee arthroplasty? 18th ESSKA Congress, Glasgow, United Kingdom	2018	1.0
Welke factoren bepalen verwachtingen van een totale knieprothese? NOV spring meeting, Amersfoort	2018	1.0
Measurement properties of the OARSI core set of performance-based measures for hip osteoarthritis. 13th European Hip Society Congress, The Hague	2018	1.0
Development and validation of a prediction model for pain and functional outcome after total knee arthroplasty using The Dutch Arthroplasty Register (LROI) data. 8th International Congress of Arthroplasty Registries, Leiden.	2019	1.0

Restklachten na een totale knieprothese kunnen worden voorspeld op basis van data verzameld voor de Landelijke Registratie Orthopedische Implantaten (LROI). DICA congres, Amsterdam.	2019	1.0
Development and validation of a prediction model for pain and functional outcome after total knee arthroplasty using The Dutch Arthroplasty Register (LROI) data. ESSKA Specialty days, Madrid.	2019	1.0
De invloed van verwachtingsmanagement op postoperatieve tevredenheid van totale knieprothese patiënten, een gerandomiseerde gecontroleerde studie: de EKSPECT-studie. NOV annual meeting, 's-Hertogenbosch.	2020	1.0

Poster presentations	**year**	**ECTS**
Measuring physical performance in hip and knee osteoarhritis. Reliability study of the recommended OARSI core set of performance-based tests. 17th ESSKA congress. Barcelona, Spain.	2016	1.0
Measuring physical performance in hip and knee osteoarhritis. Reliability study of the recommended OARSI core set of performance-based tests. 17th ESSKA congress. Barcelona, Spain.	2016	1.0
Measurement properties of the OARSI core set of performance-based measures for knee osteoarthritis. 18th ESSKA congress. Glasgow, United Kingdom.	2018	1.0
Measurement properties of the OARSI core set of performance-based measures for hip osteoarthritis. 18th ESSKA congress. Glasgow, United Kingdom.	2018	1.0
Outcome prediction for treatment of knee arthroplasty with a total knee arthroplasty. EULAR congress. Madrid, Spain	2019	1.0
Expectations of treatment result of knee osteoarthritis patients treated with a total knee arthroplasty. EULAR congress. Madrid, Spain	2019	1.0

Grant and awards	**year**
Winner 2020 van Rens Award	2020
Nominee 2019 Dutch Institute for Clinical Auditing Quality of Care Award	2019
van Rens Foundation, research grant	2018
Nominee 2018 KSSTA-Jón Karlsson Young Researcher Award in Clinical Science	2018

02. Teaching

Lecturing	year	ECTS
Performance-based testing in knee osteoarthritis. Regionale refereermiddag ROGO Zuid, Máxima MC.	2016	0.5
Expectations in Knee arthroplasty. Regionale refereermiddag ROGO Zuid, MUMC+.	2016	0.5
Klinimetrie als aanvulling op PROM's. Regionale Refereermiddag ROGO Zuid, Máxima MC.	2014	0.5

Academic teaching	year	ECTS
Tutor cluster musculoskeletal system, year 3 medical students. MUMC.	2015	1.0
Tutor cluster musculoskeletal system, year 3 medical students. MUMC.	2016	1.0

Supervising Bachelor and Masters' theses	year	ECTS
Supervising M. Mariam, physiotherapy student. Reliability of the 40m fast-paced walk test in knee osteoarthritis patients.	2015	2.0
Supervising D.A.J.M. Latijnhouwers, student human movement sciences. Measurement properties of performance-based tests in end-stage knee osteoarthritis patients.	2016	2.0
Supervising C.A. van der Velden, medical student: When is it safe to resume driving after total hip and total knee arthroplasty? a meta-analysis of literature on post-operative brake reaction times.	2017	2.0

Reviewer

2018 – present	Knee Surgery Sports Traumatology Arthroscopy
2019 – present	Arthritis Care & Research
2019 – present	EFORT open reviews

List of publications

This thesis

Tolk JJ, Janssen RPA, Haanstra TM, van der Steen MMC, Bierma Zeinstra SMA, Reijman M. The influence of Expectation modification in Knee arthroplasty on Satisfaction of PatiEnts: a randomized Controlled Trial: the EKSPECT Study. *[Submitted]*

Tolk JJ, Waarsing JH, Janssen RPA, van Steenbergen LN, Bierma-Zeinstra SMA, Reijman M. Development of preoperative prediction models for pain and functional outcome after total knee arthroplasty using The Dutch Arthroplasty Register (LROI) data. *Journal of Arthroplasty. 2020 Mar;35(3):690-698.e2.*

Tolk JJ, Janssen RPA, Haanstra TM, van der Steen MMC, Bierma Zeinstra SMA, Reijman M. Outcome Expectations of Total Knee Arthroplasty Patients: The Influence of Demographic Factors, Pain, Personality Traits, Physical and Psychological Status. *Journal of Knee Surgery*. 2019 Jul 4. [Epub ahead of print]

Tolk JJ, Janssen RPA, Prinsen CAC, Latijnhouwers DAJM, van der Steen MC, Bierma-Zeinstra SMA, Reijman M. Measurement properties of the OARSI core set of performance-based measures for hip osteoarthritis. A prospective cohort study on reliability, construct validity and responsiveness in 90 hip osteoarthritis patients. *Acta Orthopaedica*. 2019;90(1)15-20.

Tolk JJ, Janssen RPA, Haanstra TM, Bierma-Zeinstra SMA, Reijman M. Rationale and design of the EKSPECT Study: The influence of Expectation modification in Knee arthroplasty on Satisfaction of PatiEnts, a randomized Controlled Trial. *Trials*. 2018;19(1):437.

Tolk JJ, Janssen RPA, Prinsen CAC, Latijnhouwers DAJM, van der Steen MC, Bierma-Zeinstra SMA, Reijman M. The OARSI core set of performance-based measures for knee osteoarthritis is reliable but not valid and responsive. *Knee Surgery Sports Traumatology and Arthroscopy.* 2019 Sep;27(9):2898-2909.

van der Velden CA, Tolk JJ, Janssen RPA, Reijman M. When is it safe to resume driving after total hip and total knee arthroplasty? a meta-analysis of literature on post-operative brake reaction times. *Bone and Joint Journal.* 2017;99-B(5):566-576.

Tolk JJ, van der Steen MC, Janssen RPA, Reijman M. Total Knee Arthroplasty: What to Expect? A Survey of the Members of the Dutch Knee Society on Long-Term Recovery after Total Knee Arthroplasty. *Journal of Knee Surgery.* 2017;30(6):612-616.

Other Publications

Tolk JJ, van Mourik JBA. Touched by an angel. *Nederlands Tijdschrift voor Geneeskunde. 2020;164:D4419*

Tolk JJ, Haanstra TM, Reijman M. Letter to the Editor on "What Do Scottish Patients Expect of Their Total Knee Arthroplasty?". *Journal of Arthroplasty*. 2016;31(10):2374.

Tolk JJ, Willems PC, Punt I, van Rhijn LW, van Ooij A. Infection after multilevel anterior scoliosis surgery using the Cotrel-Dubousset-Hopf system. *International Journal of Spine Surgery.* 2016;10(2)

Brooijmans FAM, Tolk JJ, Janssen RPA. Het belang van ganganalyse en looptraining bij knieartrose. *Physios.* (2014) 2:38-45

Tolk JJ, Koot HWJ, Janssen RPA. Computer Navigated versus Conventional Total Knee Arthroplasty. *Journal of Knee Surgery.* 2012 Sep;25(4):347-52.

Tolk JJ, Nieuwenhuis JJ. Sesamoïditis - een klinisch perspectief. *Nederlands Tijdschrift voor Orthopaedie.* 2011;18(2)66-70

Jansen R, van Rooijen F, Toirkens JP, Besselaar AT, Tolk JJ. Paediatric Posttraumatic Cyst-Like Lesions: A Report of Two Cases and Review of the Literature. Case Reports in Orthopedics. 2020;2020:7834969